

On-the-Job Search and Optimal Schooling under Uncertainty and Irreversibility

Anna Zaharieva*

*Center for Mathematical Economics
Department of Business Administration and Economics
Bielefeld University, 33501 Bielefeld, Germany*

March 5, 2015

Abstract

This paper develops a labour market model with on-the-job search, match-specific productivity draws and an endogenous irreversible schooling decision. The choice of schooling is modeled as an optimal stopping problem which gives rise to the equilibrium heterogeneity of workers with respect to the formal education. The optimal schooling decision is characterized by the reservation productivity of students which is a monotone function of time. The reservation productivity of high ability students is increasing narrowing the range of acceptable employers in the regular labour market. This mechanism generates a positive sorting of more educated workers to more productive employers. The schooling density is downward-sloping and the equilibrium wage distribution is right-skewed with a unique interior mode. This means that the majority of workers earn wages in the middle range of the earnings distribution. At the same time there is a small proportion of employees in the beginning of their career with wages in the left tail of the earnings distribution and a small proportion of high-skilled workers earning wages in the right tail of the distribution.

JEL classification: I21, I24, J62, J64

Keywords: Optimal schooling, uncertainty, on-the-job search, wage dispersion

*E-mail: azaharieva@wiwi.uni-bielefeld.de tel.: +49-521-106-5637, fax: +49-521-106-89005. I would like to thank Daniele Checchi, Herbert Dawid, Giorgio Ferrari, Thomas Gries, Marco Leonardi and Frank Riedel for their useful comments as well as seminar participants at the University of Bielefeld and the University of Paderborn, session participants at the annual meeting of the European Society for Population Economics, annual conference of the Italian Association of Labour Economists and congress of the German Economic Association (VfS).

1 Introduction

Schooling decisions are closely linked to the situation in the labour market. Moreover, after graduating fewer workers stay with their first employer till retirement, rather most workers continue searching on-the-job and change employers in the hope of better payment. This possibility, however, may be anticipated by students leading to a feedback effect on the optimal time to stop studying. Thus the analysis of human capital accumulation will not be complete if the model does not explicitly account for the possibility of changing jobs. Some initial attempts to consider schooling as an optimal stopping problem with uncertain wage growth in the labour market have already been undertaken in the recent literature¹. However, these models consider the decision of a single student, whereas the properties of the labour market are largely underinvestigated so far. This paper attempts to fill this gap and provides a coherent framework for the analysis of the frictional labour market with endogenous schooling decisions under uncertainty, heterogeneous employers and job-to-job transitions.

Since matching is random between workers and firms some workers exit the education system earlier than others if they get an attractive job offer. Therefore, in the equilibrium workers are heterogeneous with respect to the level of schooling and the labour market is characterized by two-sided heterogeneity of workers and firms. Having investigated the equilibrium outcomes in this economy I find that the model helps to explain the following two empirical observations:

- (1). *The reservation productivity of high ability students is increasing over time reducing the range of acceptable employer types. The opposite is true for low ability workers with a decreasing reservation productivity. This mechanism generates a positive correlation between worker and firm types and explains a positive sorting of more educated workers to more productive jobs;*

This result follows directly from the decision of students to quit education. In line with the option approach to schooling, entering the labour market is associated with a positive present value of wages and is an irreversible decision. On the other hand, there is a positive value from waiting and postponing the entry. This value is attributed to the positive return from schooling as a higher level of schooling is associated with a higher expected future income. To model the labour income I follow a standard Mincer regression approach and assume that the wage is increasing in the education of the worker and the productivity of the job. Then it is possible to show that the optimal schooling decision is characterized by the reservation productivity rule saying that students should accept any job with a productivity above the reservation threshold and reject any other job. If the student never accepted any job offer and reached a maximum level of schooling such as a PhD this student may pursue an academic career. I assume that high ability workers expect to be productive in the academic market which includes positions in the field of research and teaching. This means that their reservation productivity is increasing over time narrowing the range of acceptable employers. In contrast, low ability workers are assumed to be unproductive in academic jobs and so their reservation productivity is falling over time, making them more likely to enter the regular labour market.

This mechanism implies a positive correlation between the skill level of the worker and the productivity of the job. Moreover, this finding is strongly supported by the available empirical

¹see Hogan and Walker (2007), Jacobs (2007) and Bilikic, Gries and Pilichowsky (2012)

evidence. For example, Haskel et al. (2005) consider UK manufacturing firms and report that more productive firms hire more skilled workers. Specifically, in their sample plants in the top decile of the productivity distribution hired workers with, on average, around 1/3rd of a year of extra schooling relative to plants in the bottom decile. Other empirical studies do not directly observe productivity of the job but rather use proxy variables such as a firm size, capital intensity and firm profits. In this respect, Arai (2003) finds that in Sweden a half of the positive correlation between firm profits and wages is due to systematic sorting of workers across firms. Workers with higher education and experience seem to be matched with high profit firms. Similar sorting on observables is also reported in Idson and Feaster (1990) for the US. These authors find that the firm size effect falls from 23% to only 7% when the selection of high skill workers to large firms is controlled for. More recent studies reporting sorting of high skill workers to larger more productive employers include Oi and Idson (1999), Troske (1999) and Abowd et. al. (2004).

- (2). *The equilibrium wage distribution is pick-shaped and positively skewed with a unique interior mode on the distribution support. Inexperienced unskilled workers form the left tail of the distribution, whereas experienced high skilled workers are situated in the right tail of the wage distribution. The middle range of the density is composed of the combination of inexperienced high skill workers and experienced low skill workers.*

There are two driving forces behind this result. On the one hand, the equilibrium schooling density is downward-sloping, that is the proportion of students is decreasing with a higher level of schooling. So there are less workers with a doctoral degree than those graduating from a high school. On the other hand, workers do not stay with the same employer till retirement but continue changing jobs in a search of better payment. Thus low productivity jobs are unstable and often filled with young inexperienced workers in the beginning of their career. In contrast, high productivity jobs are more stable and typically filled with more experienced older workers.

When calibrated to the US economy, the equilibrium wage density is pick-shaped with a unique interior mode on the distribution support. It is also right-skewed in accordance with the empirical evidence of the US Bureau of Labor Statistics presented in figure 1. The earnings histogram reveals that the median of the distribution is lower than the average earnings (\$43 thousands) and the distribution is positively skewed. These empirical fact is captured by the theoretical model even though the underlying productivity distribution is flat and uniform. Indeed the majority of workers in the economy have a low level of education. However these workers change jobs and move up the productivity ladder, therefore most workers earn wages in the middle range of the earnings distribution. At the same time there is a small proportion of employees in the beginning of their career with wages in the left tail of the earnings distribution and a small proportion of high-skilled workers earning wages in the right tail of the distribution. To sum up, the long right tail of the schooling density is translated into the equilibrium earnings density and therefore the model does not require unrealistically long right tails for the underlying productivity distribution (see Mortensen (2003)).

Summarizing, this paper develops a new framework for the analysis of schooling choices and labour market transitions. Moreover, the generated labour market with two-sided heterogeneity of workers and firms can replicate a number of empirical facts providing a theoretical foundation for the better understanding of empirical evidence. To the best of my knowledge there are no other studies that can explain the same empirical observations within a unified framework.

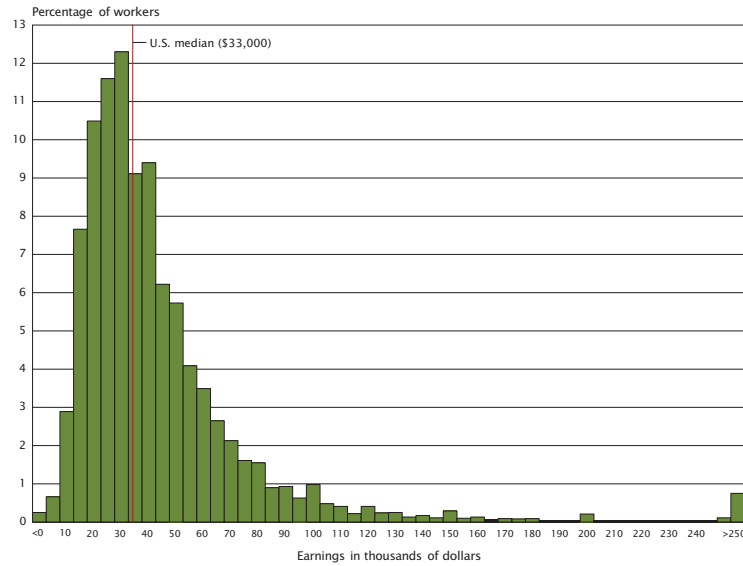


Figure 1: Annual earnings in the US. Source: Current Population Survey 2000

In addition, the model is suitable for the analysis of policy changes on schooling decisions. In particular, it turns out that more intensive job-to-job transitions discourage workers from acquiring education. This is intuitive, since more intensive on-the-job search is associated with a higher wage growth and has a positive effect on the present value of employment. Thus the possibility to stop studying and enter the labour market becomes more attractive to students, which is reflected in the accelerated exit of students from the education system. This, however, means that any policy intensifying voluntary job changes by workers will have a negative effect on the average schooling duration. One example of such policy is a commuting subsidy which is actively used in a number of European economies². According to this policy commuting costs may be deducted from the taxable income reducing the effective tax payment. In the empirical search model Rupert et al. (2009) show that workers require monetary compensation for longer commuting times. Specifically, they find that a salary increase of 28.5% is required to compensate workers for one hour longer commuting per day. Similar results are also reported by Ommeren et al. (2000), who find that the average willingness to pay for one hour commuting per day is estimated to equal almost half of the hourly wage. Thus it can be expected that a more generous commuting subsidy will reduce the cost of accepting distant jobs and increase the intensity of job-to-job transitions in the economy. This paper shows that one potential adverse effect of this policy is a lower attractiveness of education and a reduction of schooling duration.

Related literature on optimal schooling under uncertainty can be generally divided into two research directions. First, Williams (1979) suggests to model education as a risky investment and follows a portfolio approach. The focus of his paper is on the optimal allocation of time invested in education over the life cycle of the worker. This paper proves the optimal separation between allocations to education and investments in risky assets and justifies the separate treatments of education and risky assets in labor economics and finance. The approach of Williams (1979) is

²For example, in Germany workers obtain a subsidy of 30 cents per km with a maximum payment of 4500 EUR per year. Denmark requires a minimum distance traveled of 12 kilometers and then grants an allowance of 15 cents. Luxembourg allows a standard deduction of 396 EUR for distances of less than 4 kilometers which rises to a maximum of 2970 EUR.

extended by Hanchane et al. (2006). In this paper the authors introduce a positive correlation between the process of human capital accumulation and the value of education in the labour market which can be interpreted as an ability bias. This correlation leads them to the conclusion that a higher wage risk has a positive effect on the investment in human capital. This is in contrast to the original finding of Williams (1979), Eaton and Rosen (1980) and even earlier by Levhari and Weiss (1974) in a two-period model.

The second approach was originated by the contribution of Hogan and Walker (2007) who were the first to apply the theory of real options to the human capital investment decision. The main idea of this approach is to separate the two components of human capital: formal education and experience. The formal education is then modeled as an irreversible decision so the study by Hogan and Walker (2007) is most closely related to the present paper. However, there are a number of significant deviations. First, this paper recognizes that job offers and therefore the possibilities to enter the labour market are rare events, so the process of job arrivals is modeled by means of a Poisson process. Second, there is a maximum duration of schooling in the present study such as the doctoral degree. This modification is not only a realistic feature of the education system, but it also generates the dynamics of the reservation productivity and wage which is absent in the original study by Hogan and Walker (2007).

Next, the growth rate of income is not a constant in the present study, rather workers continue searching on-the-job and the growth rate of their income is stochastic. Similar to the present study Bilkic et al. (2012) also relax the assumption of a constant growth rate and assume that the individual income dynamics is described by the geometric Brownian motion. As a result in their model the entry threshold is necessarily increasing over time, that is, the student wants to be compensated for the additional education period and costs by a higher entry-level income. In the present study the reason for the dynamics of the entry threshold is principally different, it is the remaining time to the maximum duration of schooling which makes students more/less choosy over time depending on their productivity in the academic market (which is used as a proxy for their ability).

Finally, the most important difference of the present study and the models of Hogan and Walker (2007) and Bilkic et al. (2012) is a labour market perspective of the analysis rather than a problem of a single agent. In this respect the paper can be viewed as an extension of a model by Burdett and Mortensen (1998) to the population of workers with heterogeneous skills. In the field of search and matching theory the study which is closest to the present one is a recent paper by Burdett et al. (2011). Human capital in their model takes the form of a general experience of the worker and is accumulated over time (learning-by-doing). Similar to the present study their model can replicate the hump-shaped distribution of wages with a "fat" right tail. However, skill heterogeneity is exogenous in their model, so they can not reproduce the positive sorting of more educated workers to more productive jobs. Similar search models with learning-by-doing include Bagger et. al (2011) and Yamaguchi (2010).

Other studies on the education choice under uncertainty include Altonji (1991), Groot and Oosterbeek (1992) and Jacobs (2007). Altonji (1991) considers the risk for the student of not being able to complete the education program, so his model can explain the endogenous drop-out rate from the college. This realistic feature of the model is also included in the present study. Groot and Oosterbeek (1992) analyze the effect of the unemployment risk for the optimal level

of schooling. Their model can replicate the fact that unemployment is a decreasing function of schooling. Finally Jacobs (2007) investigates the option for the young individual to start education or postpone it to the later period of time. The focus of his paper is primarily on the effect of taxation on schooling.

The plan of the paper is as follows. Section 2 explains notation and the general economic environment. Section 3 presents the problem of optimal stopping and gives a full characterization of the individual decision to enter the labour market. Section 4 describes the properties of the labour market when workers are heterogeneous with respect to their schooling level. Section 5 incorporates the risk of unemployment which allows me to calibrate the model to the US labor market. Section 6 concludes the paper.

2 Labour market modeling framework

The labour market is characterized by the following properties. Time is continuous and at every point in time $n(0)$ new agents enter the education system forming a pool of students. Similarly ρ is a permanent exit rate of all agents into non-participation, so the total measure of participating agents in the economy is a constant. Students are ex-ante identical upon entry, the time horizon is infinite and the discount rate is r . The flow cost of a student is denoted by y , reflecting the costs of accommodation and tuition fees. At the same time the flow income of a student is denoted by z . This can be a scholarship or a flow salary from a student job. Variable T reflects the maximum duration of schooling. Students with a maximum level of schooling T enter the academic market with an exogenous present value of income $V(T) = a/(r + \rho)$. The academic market includes positions in research and teaching. This market is not explicitly modeled in the paper but it has an impact on the decision to exit the education system. In the following I am going to perform comparative statics with respect to a and analyze its impact on the students' decision to stop studying. Thus differences in a can be intuitively interpreted as differences in the in-born abilities of students. For example, high ability students would be more productive in the academic job, which is associated with a higher value of a . In contrast, low ability students will be relatively unproductive in the academic job, which is associated with a low value of a .

In the process of acquiring education students are randomly matched with firms and obtain job offers at the Poisson arrival rate λ . If the job offer is accepted the student enters the pool of employees with an acquired level of schooling. This is an irreversible decision. Let w be the flow wage of the employee which is given by the extended Mincer wage regression of the form:

$$\ln w = \alpha + \beta s + x + \varepsilon$$

where α is a constant term, β is the return to schooling, s is the acquired level of schooling, x is a firm fixed effect and ε is the random luck component. The linear relationship between the log-wage and the schooling level is supported on the empirical level, for example, Park (1994) reports that the linear functional form provides a good fit to the data and Card (1999) concludes that it is reasonable "to assign an estimate of the years of completed education to each reported education class and assume a linear functional form" (page 1808). From the perspective of workers, there are two sources of wage uncertainty, on the one hand, firms are heterogeneous

and workers don't know in advance the type of the employer they will be matched with, i.e. a high or low realization of x . The firm fixed effect, for example, can be interpreted as a size of the employer. There is strong empirical evidence that larger firms pay higher wages after controlling for the observed worker characteristics, see Oi and Idson (1999) and Troske (1999). On the other hand, there is a random luck component ε reflecting the quality of the match between the firm and the worker. To simplify the representation of the model, let $p \equiv e^{\alpha+x+\varepsilon}$ denote the match-specific productivity of the job. Thus the productivity can be high because the firm is large and productive, i.e. a high realization of x , or because the worker is a good match for the job, i.e. a high realization of ε . So the wage equation becomes: $w = pe^{\beta s}$ reflecting the fact that wages are increasing in the schooling level s and the match-specific productivity p . Let $F(p)$ denote the distribution of match-specific productivities $p \in [0, \bar{p}]$. Workers learn the realization of p upon the match with the employer and have to decide if they accept the job or not. Employed workers continue searching on-the-job and are randomly matched with new employers at rate λ_e .

Finally, in sections 3 and 4 the risk of becoming unemployed is not considered to simplify the representation. Nevertheless, unemployment is included into the model in section 5 to make the framework more realistic. This extended model is then calibrated to match parameters in the US labour market, however it should be noted that the state of unemployment does not change the qualitative predictions of the model.

3 Bellman equations and optimal stopping

This section considers the present values of students and workers in the economy. Let $V(t)$ denote the expected present value of future income of a student at time t , which means that the accumulated level of schooling of this student is equal to t . Similarly, denote $W(t, p)$ – the expected present value of a worker employed in a job with a match-specific productivity p and schooling t . The Bellman equation for $V(t)$ is given by (for the derivation see Appendix I):

$$(r + \rho)V(t) = (z - y)(1 - e^{-\lambda(T-t)}) + \lambda \int_0^{\bar{p}} \max(W(t, p) - V(t), 0) dF(p) + \dot{V}(t) \quad (3.1)$$

The first term in this equation denotes the present value of the net student income $z - y$ over the remaining period of time $T - t$. The second term is the option value of waiting and postponing the entry into the labour market. Here the gain from waiting is twofold. On the one hand, students can wait for more productive jobs in the future, and on the other hand, waiting directly leads to a higher level of schooling. Both of these factors imply a higher expected future wage $w = pe^{\beta t}$. Next denote $k = r + \rho$ and let s be the moment of time when the student enters the labour market. Therefore s is the acquired level of schooling which remains constant after the market entry. Bellman equation for $W(s, p)$ is given by:

$$\begin{aligned} kW(s, p) &= pe^{\beta s} + \lambda_e \int_0^{\bar{p}} \max(W(s, x) - W(s, p), 0) dF(x) \\ &= pe^{\beta s} + \lambda_e \int_p^{\bar{p}} (1 - F(x)) W'_p dx \end{aligned} \quad (3.2)$$

This equation shows that workers employed with a productivity p accept job offers from new employers only if their productivity level exceeds p . Next, differentiate $W(s, p)$ with respect to p to obtain $W'_p = e^{\beta s} / (k + \lambda_e(1 - F(p)))$. So the present value of employment is separable in productivity and schooling: $W(s, p) = \Omega(p)e^{\beta s}$ where the auxiliary variable $\Omega(p)$ is given by:

$$k\Omega(p) = p + \lambda_e \int_p^{\bar{p}} \frac{(1 - F(x))dx}{k + \lambda_e(1 - F(x))} \quad (3.3)$$

Note that $W(s, p)$ is increasing in both arguments. In addition, in the absence of on-the-job search (when $\lambda_e=0$) variable $W(s, p)$ is simplified to yield $pe^{\beta s}/k$. This means that the transformed present value of employment is a steeper function of productivity in the absence of on-the-job search. This is explained by the fact that the option value of waiting for better paid jobs falls with a higher p . Further consider the decision of students to quit schooling, which is an irreversible decision. Building upon the separation result from above insert $W(s, p) = \Omega(p)e^{\beta s}$ into the present value equation for $V(t)$ and consider the simplified case $z = y$, which means that the costs and expenses from schooling (tuition fees) are fully covered by the student scholarship:

$$kV(t)e^{-\beta t} = \lambda \int_0^{\bar{p}} \max(\Omega(p) - V(t)e^{-\beta t}, 0) dF(p) + \dot{V}(t)e^{-\beta t}$$

Denote $\phi(t) = V(t)e^{-\beta t}$, so that $\dot{\phi}(t) = \dot{V}(t)e^{-\beta t} - \beta\phi(t)$. In addition, let $p_r(t)$ be the reservation productivity of students which is defined as $\Omega(p_r(t)) = \phi(t)$. Intuitively, at the productivity $p_r(t)$ the student is indifferent between accepting the job and obtaining the present value $\Omega(p_r(t))e^{\beta t}$ or rejecting the job and continuing education yielding the present value $V(t)$. Bellman equation for the auxiliary variable $\phi(t)$ reads as:

$$(k - \beta)\phi(t) = \lambda \int_{p_r(t)}^{\bar{p}} (\Omega(p) - \phi(t)) dF(p) + \dot{\phi}(t)$$

First, consider the case when the reservation productivity is constant over time so that $\dot{\phi}(t) = 0$ and $\Omega(p_r) = \phi$. This is a special case which applies if and only if $p_r = p_r(T)$ where the end-point productivity $p_r(T)$ can be obtained from the present value expression in the academic market: $\Omega(p_r(T)) = \phi(T) = V(T)e^{-\beta T}$. For job offers with $p > p_r(T)$ students with a maximum level of education T prefer to work in the regular labour market. On the contrary, if the market job offer is $p < p_r(T)$ these students pursue a career in research and teaching. Further observe that the benchmark productivity p_r indicates the fixed point in the dynamics of $p_r(t)$ and is implicitly given by the following equation:

$$(k - \beta)\Omega(p_r) = \lambda \int_{p_r}^{\bar{p}} (1 - F(p))\Omega'_p dp \equiv \Phi(p_r) \quad (3.4)$$

Let $k > \beta$, the transformed present value of employment $\Omega(p_r)$ is increasing in p_r in the range $[0, \bar{p}]$, while $\Phi(p_r)$ is decreasing since $\Phi'(p_r) = -\lambda\Omega'(p_r)(1 - F(p_r)) < 0$. Moreover, it is also true that $\Phi(\bar{p}) = 0$. This means that equation (3.4) defines a unique positive value p_r if $(k - \beta)\Omega(0) < \Phi(0)$ which is equivalent to $(k - \beta)\lambda_e < k\lambda$. In the following assume that this condition is satisfied. Also note that a higher return to schooling β shifts the left-hand side of equation (3.4) downwards which has a positive effect on the benchmark productivity p_r . Therefore the

duration of schooling is generally increased and students remain longer in the education system if β is higher.

Next, recall that a denotes the fixed flow income in the academic market (which is a proxy for student's ability), so that $V(T) = a/k$ and $\Omega(p_r(T)) = ae^{-\beta T}/k$ which yields the end-point boundary $p_r(T)$. This means that the optimal student reservation productivity $p_r(t)$ is constant and coincides with p_r if and only if the flow income in the academic market is equal to $a_r \equiv k\Omega(p_r)e^{\beta T}$. Otherwise $p_r(t)$ is given by the following first order nonlinear ordinary differential equation³:

$$\Omega'_p \dot{p}_r(t) = (k - \beta)\Omega(p_r(t)) - \Phi(p_r(t)) \quad (3.5)$$

This equation is autonomous since the right-hand side does not explicitly depend upon the time variable. The discretized form of the differential equation (3.5) is illustrated in figure 2. The corresponding curve for $\dot{p}_r(t)$ is cutting the 45-degree line at the stationary point p_r . Therefore if the end-point $p_r(T)$ is situated above p_r then the reservation productivity function $p_r(t)$ will be increasing over time with a starting point $p_r(0)$ situated above p_r . Otherwise it will be decreasing, hence it follows that $p_r(t)$ is a monotone function of time. Solution of the differential equation (3.5) is implicitly given by the following function:

$$\int_C^{p_r(t)} \frac{\Omega'_p(x)dx}{(k - \beta)\Omega(x) - \Phi(x)} = t, \quad t = 0..T \quad (3.6)$$

where the constant term $C = p_r(0)$ can be found by inserting $t = T$. This is the initial condition for the differential equation (3.5), hence the optimal student reservation productivity $p_r(t)$ can be obtained for every t . Its properties are described in proposition 1.

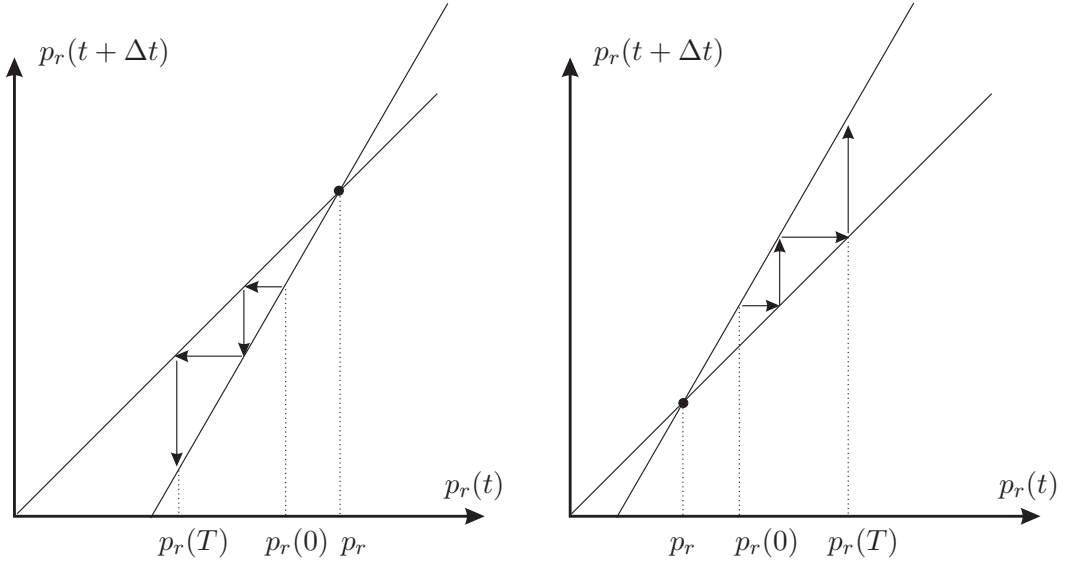


Figure 2: Dynamics of the student reservation productivity

³For a more detailed treatment of non-stationarity in search theory see Van den Berg (1990)

Proposition 1: *The student reservation productivity $p_r(t)$ is increasing for every $t < T$ and $p_r < p_r(t) < p_r(T)$ if $a > a_r = k\Omega(p_r)e^{\beta T}$. Otherwise $p_r > p_r(t) > p_r(T)$ and $p_r(t)$ is decreasing for every $t < T$. The student reservation wage $w_r(t) = p_r(t)e^{\beta t}$ is an increasing function of t if $\dot{p}_r(t) \geq 0$.*

Proof: Appendix II.

As can be seen from proposition 1, the reservation productivity of high ability students ($a > a_r$) is increasing over time. This is because high ability students are sufficiently productive in the academic market. When students enter the education system, future income $V(T)$ is heavily discounted over the maximum schooling time T and is multiplied by the probability of not entering the labour market. So the reservation productivity is low initially but then rises as the valuable future perspective to enter the academic market comes nearer and becomes more likely. This effect is reversed if the academic perspective is not attractive, i.e. $a < a_r$, which happens for students with low abilities. Further, consider the role of on-the-job search. In particular, there is a question whether efficiency of on-the-job search can reverse dynamics of the reservation productivity. Proposition 2 is summarizing the result.

Proposition 2: *Let p_r and p_0 denote the benchmark productivities with and without on-the-job search respectively. In addition, let $a_r = k\Omega(p_r, \lambda_e)e^{\beta T}$ and $a_0 = k\Omega(p_0, 0)e^{\beta T}$ be the corresponding flow incomes in the academic market. Then $p_0 > p_r$ but $a_0 < a_r$. This means there exists an interval of flow values $a \in (a_0, a_r)$ such that the reservation productivity of students is decreasing in the economy with on-the-job search but it is increasing in the economy without on-the-job search: $\dot{p}_r(t) < 0$, $\dot{p}_0(t) > 0$.*

Proof: Appendix III.

Note that economy without on-the-job search can be characterized by considering the implications of a lower job-finding rate λ_e (with a corner case $\lambda_e = 0$). If λ_e falls there is a twofold effect on the transformed present value of employment $\Omega(p_r, \lambda_e)$. First, from equation (3.2) it immediately follows that employed workers are worse off if the probability of finding a better paid job is reduced. This is a negative effect. On the other hand, a lower job-finding rate λ_e makes students more choosy, so the benchmark reservation productivity p_r is increased. This is due to the fact that the option of waiting for better paid jobs becomes more valuable as the probability of meeting new employers on-the-job is reduced. Proposition 2 shows that the first effect is dominating and a lower λ_e has a negative effect on $\Omega(p_r, \lambda_e)$, so the benchmark value $a_0 = k\Omega(p_0, 0)e^{\beta T}$ in the economy without on-the-job search is lower than a_r . Intuitively, the present value of income is higher in the economy with on-the-job search. Therefore, students with abilities a in the range (a_0, a_r) are estimating their income in the academic market as "relatively low", so their reservation productivity is falling over time in the economy with on-the-job search. This is different when the option to search on-the-job is not available and their income in the academic market is estimated as "relatively high" compared to the regular market.

3.1 Special case: uniform productivity distribution

This section investigates the model with a uniform productivity distribution $F(p) = p$, $p \in [0, 1]$ so the maximum match-specific productivity is normalized to 1. From the perspective of students this distribution is associated with strong uncertainty about the future job as high and low productivity jobs are equally likely to be drawn. With on-the-job search the benchmark

productivity p_r is implicitly given by (see Appendix IV):

$$p_r = \left(1 - \frac{\lambda_e(k - \beta)}{\lambda k}\right) \left[1 + \frac{k}{\lambda_e} \ln \frac{k}{k + \lambda_e(1 - p_r)}\right]$$

The benchmark threshold p_r for entering the labour market is obtained at the intersection between the upward-sloping curve $(k - \beta)\Omega(p_r)$ and the downward-sloping curve $\Phi(p_r)$. Recall from the previous subsection that $\lambda_e(k - \beta) < \lambda k$ and therefore both multipliers in the above equation are less than 1. So there exists a unique $p_r \in [0..1]$. Next consider an economy without on-the-job search. The transformed present value of employment is then linear in the productivity of the job: $\Omega(p) = p/k$ and the option value of waiting is a quadratic function $\Phi(p_0) = 0.5\lambda(1 - p_0)^2/k$. Therefore the benchmark productivity p_0 is equal to the lower root of the quadratic polynomial $(k - \beta)p_0 = 0.5\lambda(1 - p_0)^2$:

$$p_0 = 1 - \frac{k - \beta}{\lambda} \left(\sqrt{1 + \frac{2\lambda}{k - \beta}} - 1 \right) < 1$$

Denote the second root of this polynomial by $p_1 > 1$. Clearly, a higher return to schooling β leads to higher benchmark reservation productivities p_r and p_0 . This means that a higher β may reverse dynamics of the reservation curves $p_0(t)$ and $p_r(t)$ from increasing to decreasing over time. Intuitively, if the return to schooling is low a large fraction of students exits the education system in the beginning of their studies (if the reservation productivity curve is positively sloped). In this case a higher value of β may change the slope of the reservation curve and reduce the proportion of students exiting the education system with a low level of schooling. Consequently, a higher value of β leads to a higher average level of schooling in the population.

In the absence of on-the-job search the reservation function $p_0(t)$ is explicitly given by:

$$p_0(t) = \frac{p_1(C_0 - p_0) - p_0(C_0 - p_1)e^{-0.5\lambda(p_1 - p_0)t}}{C_0 - p_0 - (C_0 - p_1)e^{-0.5\lambda(p_1 - p_0)t}}$$

See appendix IV for the derivation. Note that this equation implies $p_0(0) = C_0$ so that in accordance with proposition 1 $p_0(t)$ is an increasing function of time if $C_0 = p_0(0) > p_0$ and a decreasing function of time otherwise. Both reservation productivities $p_r(t)$ and $p_0(t)$ are illustrated in figure 3. The model parameters are chosen to calibrate the full model to the semi-annual data in the United States. The calibration is presented in details in section 5 and the chosen parameters are: $r = 0.025$, $\rho = 0.022$, $\beta = 0.037$, $\lambda = 0,181$ and $\lambda_e = 0.341$. In addition, the maximum level of schooling T is set to 24 semesters of voluntary education (12 years), including high school, bachelor, master and doctoral degrees.

For the above parameter choices the benchmark reservation productivities are given by: $p_0 = 0.72$, $p_r = 0.47$ with the corresponding benchmark abilities (wages in the academic market) $a_0 = 1.76$ and $a_r = 1.92$. Let $\Delta a = a_r - a_0 = 0.16$. The left panel of figure 3 presents reservation productivity curves $p_r(t)$ and $p_0(t)$ for the three ability values: $a = [a_r - \Delta a, a_r, a_r + \Delta a]$. First, at every point in time the reservation productivity of students in the economy with on-the-job search is lower compared to the economy without on-the-job search. This means that students are less choosy about job offers if there is a possibility to quit jobs for better paid ones.

Second, in line with a theoretical prediction the reservation curve $p_0(t)$ is upward-sloping for

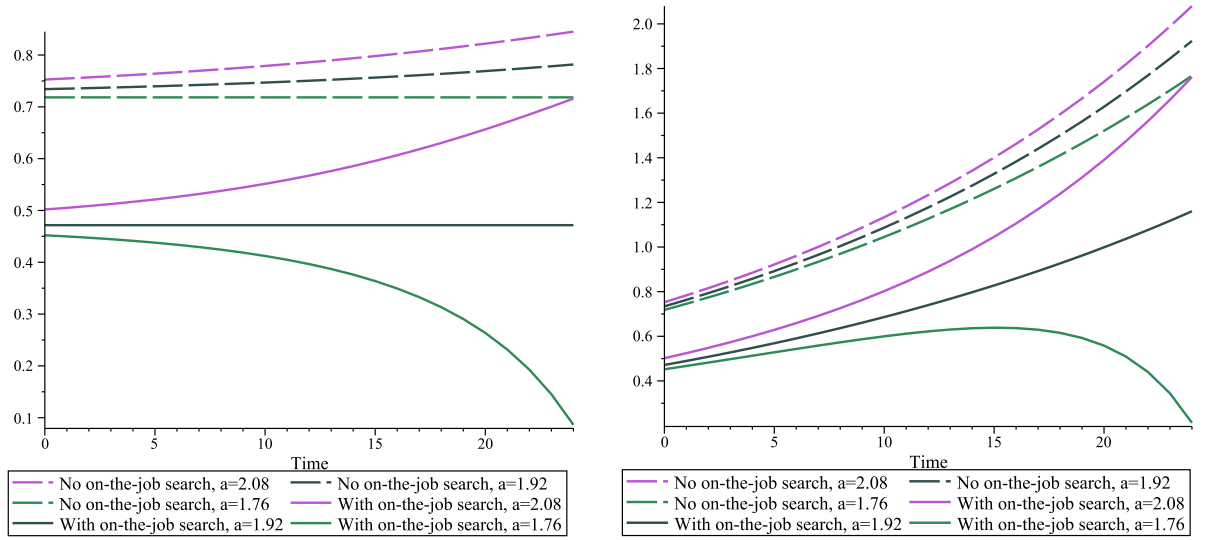


Figure 3: Left panel: reservation productivities. Right panel: reservation wages

students with intermediate abilities $a_0 < a < a_r$ while the curve $p_r(t)$ is downward-sloping in this range. So these students become more choosy about job offers over time in the economy without on-the-job search while the reverse applies in the economy with on-the-job search. Moreover, the variation of reservation productivities with a is stronger in the economy with on-the-job search. In fact the reservation productivity curves are steeper over time the higher the intensity of on-the-job search λ_e . To see this consider the maximum level of schooling $s = T$ with the corresponding reservation productivity $p_r(T)$. The change in $p_r(T)$ given a higher terminal wage a is given by:

$$k \frac{\partial p_r(T)}{\partial a} = (k + \lambda_e(1 - F(p_r(T))))e^{-\beta T}$$

Therefore, a higher on-the-job search intensity λ_e makes the terminal reservation productivity $p_r(T)$ more sensitive to changes in a . Moreover this initial effect is further reinforced by a lower level of $p_r(T)$. This higher variation in the terminal reservation productivity is then translated into the whole curve $p_r(t)$ making it steeper over time. Finally, note that the reservation wage $w_r(t) = p_r(t)e^{\beta t}$ is not necessarily a monotone function of time. This is illustrated on the right panel of figure 3. For $a = a_0$ the reservation wage in the economy with on-the-job search has an internal maximum at $t = 14$ semesters. This is explained by the fact that the reservation productivity curve is flatter for low levels of schooling since the terminal wage in the academic market is heavily discounted over the remaining period of time. So the positive effect of higher schooling on the reservation wage is dominating. In contrast when the remaining schooling time $T - t$ is low the slope of $p_r(t)$ is changing faster so the negative effect of a lower reservation productivity is dominating at high levels of schooling.

Finally, note that all of the above results are derived for the case when the exit from education is considered to be an irreversible decision. Even though there is empirical support of this assumption, in every dataset there is a small proportion of students returning to the education system after a delay of more than one year. For example, Rowan-Kenyon (2007) considers high school graduates from 1992 in the United States and follows their activities in the next 8 years.

The author finds that 68% of graduates have immediately enrolled in college without delay, about 9% of graduates enrolled within the next year, 14% left the education system and never returned. And about 9% of students returned to the education system after a delay of more than 1 year. In a similar sample Johnson (2013) finds that 53% of high school graduates enrolled directly in college, 30% stopped schooling and never returned to the education system, and only 17% of students continued studying after some delay. This evidence confirms that among high school graduates most students either continue their studies directly or start working and do not return into the education system. So I keep the assumption of irreversibility till the end of the paper. Nevertheless, appendix V presents an overview of the model where workers may continue studying if their employment in the regular market was not successful. I find that the reservation productivity is then adjusted to account for the possibility of losing the job and continuing the education.

4 Aggregate variables

This section deals with the equilibrium distributions of schooling, productivities and wages. Every student optimally exits the education system upon receiving a job offer with a productivity $p > p_r(t)$, an event which happens at rate $\rho + \lambda(1 - F(p_r(t)))$. As a result workers in the labour market are endogenously heterogeneous with respect to the accomplished education level s . Let $n(s)$ denote the total measure of students with a schooling level s :

$$\dot{n}(s) = -(\rho + \lambda(1 - F(p_r(s))))n(s) \quad (4.7)$$

where $n(0)$ is the total inflow of students into the education system per unit time. From the above differential equation, the total measure of students $n(s)$ is given by:

$$\ln n(s) = \ln n(0) - \rho s - \lambda \int_0^s (1 - F(p_r(t)))dt$$

Denote $g(s)$ – the probability density function of schooling types, so that $g(s) = n(s)/n$, where n is a total measure of students in the economy: $n = \int_0^T n(t)dt$. From equation (4.7) it follows that $g(s)$ is a decreasing function of s , that is the proportion of students is decreasing with a higher level of schooling. So there are less students pursuing a PhD degree than those graduating from a high school. Further, let $E(p|s)$ be the total measure of workers employed with a productivity below or equal to p and a fixed schooling level s . Students $n(s)$ enter the group $E(p|s)$ at rate $\lambda(F(p) - F(p_r(s)))$, which is their primary employment. In addition, employed workers may retire at rate ρ or quit the job for a better paid one at rate $\lambda_e(1 - F(p))$, this means:

$$\dot{E}(p|s) = \lambda(F(p) - F(p_r(s)))n(s) - (\rho + \lambda_e(1 - F(p)))E(p|s)$$

Let $E(s) \equiv E(\bar{p}|s)$ denote the total measure of employees with a schooling level s and $E \equiv \int_0^T E(s)ds$ – the total employment in the labour market. In the stationary equilibrium the inflow of workers into the group $E(p|s)$ should be equal to the outflow of workers from this group, so that $\dot{E}(p|s) = 0$. This condition should also hold for the aggregate employment $\dot{E} = 0$, so that $\rho E = \lambda n \int_0^T (1 - F(p_r(s)))g(s)ds$, where the right-hand side is an average student job-finding

rate multiplied by the total number of students n . Intuitively, a higher job arrival rate λ , a higher number of students n , and a higher average job acceptance rate contribute to the higher stock of employment E . Differentiating $E(p|s)$ with respect to p yields the joint probability density function $q(p, s) = h(p|s) \cdot k(s)$, where $h(p|s) = [\partial E(p|s)/\partial p]/E(s)$ and $k(s) = E(s)/E$:

$$q(p, s) = \underbrace{\frac{\rho f(p)(\rho + \lambda_e(1 - F(p_r(s))))}{(\rho + \lambda_e(1 - F(p)))^2(1 - F(p_r(s)))}}_{h(p|s)} \cdot \underbrace{\frac{(1 - F(p_r(s)))g(s)}{\int_0^T (1 - F(p_r(s)))g(s)ds}}_{k(s)} \quad (4.8)$$

The first multiplier in this equation $h(p|s)$ is a productivity density function in the population of workers with s years of schooling. To simplify the analysis consider a hypothetical case $p_r(s) = 0 \forall s$, which means that the support of the distribution remains unchanged in the equilibrium $[0, \bar{p}]$. The productivity density function is then given by $\rho f(p)(\rho + \lambda_e)/(\rho + \lambda_e(1 - F(p)))^2$ in the economy with on-the-job search and $f(p)$ otherwise. On-the-job search allows workers to climb a wage and productivity ladder, so in the equilibrium there are disproportionately less workers employed at low productivities $h(0) = \rho f(0)/(\rho + \lambda_e) < f(0)$ and disproportionately more workers employed at high productivities $h(\bar{p}) = f(\bar{p})(\rho + \lambda_e)/\rho > f(\bar{p})$. Therefore, on-the-job search has a positive primary effect on the average productivity of workers if the distribution support remains unchanged. Next, consider a more general case when $p_r(s) > 0$ and recall from section 3 that the reservation productivity is higher in the economy without on-the-job search. If workers anticipate a possibility to change jobs in the future the range of accepted productivity values is enlarged. This is a negative secondary effect of on-the-job search on the average productivity. To conclude this part, a more flexible labour market does not necessarily lead to a higher average productivity for comparable workers (with the same schooling s).

Now compare groups of workers with different years of schooling. If the reservation productivity is increasing over time more educated workers reject a larger number of applications. So the range of accepted productivities is narrowing with a higher level of schooling and the average productivity is increasing. This means that more educated workers on average are employed in more productive firms, so there is positive assortative matching on observable worker and firm characteristics. This prediction is supported by the empirical evidence, for example Oi and Idson (1999), Troske (1999) and Abowd et. al. (2004) report a positive sorting of high skill workers into larger firms. Specifically, these studies show that larger more productive firms pay higher wages. Moreover, this wage premium is reduced but not eliminated by the positive sorting of workers. This means that sorting of more educated workers to larger firms explains a part of the firm size premium, but it still exists after controlling for the observable worker characteristics.

The second multiplier $k(s)$ in the equation for $q(p, s)$ is an equilibrium density function of schooling types in the population of workers. If the academic market is sufficiently attractive and the reservation productivity $p_r(s)$ is increasing over time, then the student job acceptance rate $(1 - F(p_r(s)))$ is above the average for low values of schooling and it is below the average for higher levels of schooling. Consequently, there are disproportionately more employees with low schooling than students (i.e $k(0) > g(0)$ and $k(T) < g(T)$). The opposite holds when the reservation productivity is decreasing over time, in addition $k(s) = g(s) \forall s$ if the reservation productivity is constant.

Finally, an equilibrium distribution of earnings $Q(w)$ can be obtained by integrating the joint

density $q(s, p)$ over productivity and schooling: $Q(w) = P\{pe^{\beta s} < w\}$. Note that individual wages are increasing in schooling s and the match-specific productivity p . If in addition the reservation productivity $p_r(t)$ is increasing over time then the lowest wage in the economy is equal to $p_r(0)$. This wage is paid to unskilled workers with zero voluntary schooling employed in jobs with a lowest productivity. In contrast, the highest wage in the economy is $\bar{p}e^{\beta T}$, which is paid to workers with a highest level of schooling T employed in most productive jobs.

The equilibrium distribution of wages is then given by:

$$Q(w) = \begin{cases} Q_1(w)/E & \text{if } p_r(0) < w < \bar{p} \\ Q_2(w)/E & \text{if } \bar{p} < w < \bar{p}e^{\beta T} \end{cases} \quad \text{where} \quad Q_1(w) = \int_0^T \max[0, E(we^{-\beta s}|s)] ds$$

$$\text{and} \quad Q_2(w) = \frac{\lambda n}{\rho} \int_0^{\frac{1}{\beta} \ln(\frac{w}{\bar{p}})} (1 - F(p_r(s))) g(s) ds + \int_{\frac{1}{\beta} \ln(\frac{w}{\bar{p}})}^T \max[0, E(we^{-\beta s}|s)] ds$$

Function $Q(w)$ satisfies the standard properties, since $Q_1(p_r(0)) = 0$, $Q_1(\bar{p}) = Q_2(\bar{p})$ and $Q_2(\bar{p}e^{\beta T}) = 1$. However, the earnings density is discontinuous at the point $w = \bar{p}$. The following section shows that the point $w = \bar{p}$ delivers an internal maximum of the equilibrium earnings distribution, so the respective density function is pick-shaped with a unique internal mode on the distribution support.

4.1 Special case: uniform productivity distribution

This section continues analysis of the labour market with a uniform productivity distribution. For the special case when the reservation productivity is a constant, the schooling density coincides with the truncated exponential distribution:

$$g(s) = \frac{(\rho + \lambda(1 - F(p_r)))e^{-(\rho + \lambda(1 - F(p_r)))s}}{1 - e^{-(\rho + \lambda(1 - F(p_r)))T}} \quad s = 0 \dots T$$

Figure 4 illustrates probability density functions of schooling types in the pool of students. These density functions are presented for the three respective ability values: $a = [a_r - \Delta a, a_r, a_r + \Delta a]$ with and without on-the-job search. All density functions are decreasing which implies that the proportion of students is decreasing in s . Let $g_0(s)$ denote the corresponding schooling density in the economy without on-the-job search. For the case of the uniform productivity distribution this function can be obtained explicitly as (see Appendix IV):

$$g_0(s) = \frac{n(0)}{n(p_1 - p_0)^2} \left[C_0 - p_0 - (C_0 - p_1)e^{-0.5\lambda(p_1 - p_0)s} \right]^2 e^{-(\rho + \lambda(1 - p_1))s}$$

If the reservation productivity $p_0(t)$ is constant over time, it holds that $C_0 = p_0$, so the above density function simplifies to yield $g_0(s) = n(0)e^{-(\rho + \lambda(1 - p_0))s}/n$.

Next consider respective proportions of unskilled workers ($g(0)$) in both economies, which can be obtained as $g(0) = n(0)/n$. If the entry of workers $n(0)$ is the same, then differences in the proportions of unskilled workers are solely attributed to differences in the total numbers of students n . Therefore, comparing the numbers of students reveals the following:

$$n(0) \int_0^T e^{-(\rho + \lambda)s + \lambda \int_0^s F(p_r(t)) dt} ds < n(0) \int_0^T e^{-(\rho + \lambda)s + \lambda \int_0^s F(p_0(t)) dt} ds$$

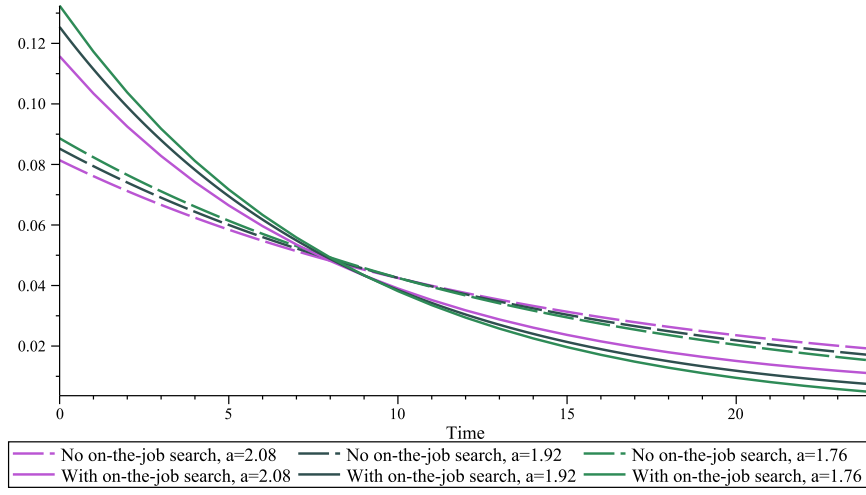


Figure 4: Probability density functions of schooling types $g(s)$

so the total number of students is lower in the economy with on-the-job search due to the lower reservation productivity curve $\int_0^s F(p_r(t))dt < \int_0^s F(p_0(t))dt$. Consequently, the model predicts that on-the-job search improves the outside opportunities of workers and unambiguously leads to a lower number of students n and a higher proportion of unskilled workers $g(0)$. In addition, this finding implies that the schooling density function is steeper in the economy with on-the-job search and the average level of schooling is lower. These predictions are confirmed in the numerical example. For the benchmark case $a = a_r = 1.92$ the total numbers of students are given by $7.97n(0)$ and $11.73n(0)$ with and without on-the-job search respectively. Similarly, the average levels of schooling are equal to 7.0 and 8.9 semesters in the two economies. These findings have important policy implications. For example, if policy is designed in a way intensifying on-the-job search, it may have an adverse effect on the number of students and the average schooling duration. As mentioned in the introduction, one example of such policy is a commuting subsidy, which is actively used in a number of European economies. A more generous commuting subsidy may increase the intensity of job-to-job transitions as workers will be more likely to accept distant jobs. In this situation the model predicts that the labour market will become more attractive to students which will be associated with an accelerated exit of students from education.

Differences in students' abilities reflected in the attractiveness of the academic market a are also illustrated on figure 4. Here both models predict that more able students have higher reservation productivities and are less likely to exit the education system. This leads to flattening of the schooling density functions, so the average duration of schooling is unambiguously increasing with the students' ability a . For example, in the model without on-the-job search the average schooling is raised to 9.2 semesters when a is increased to 2.08, while it falls down to only 8.6 semesters when a is reduced to 1.76.

Next consider the equilibrium productivity distributions $h(p|s)$ conditional on the level of schooling. These are illustrated in figure 5 where the chosen schooling levels are equal to $s = 0$ (the left panel) and $s = 24$ (the right panel). Note that the productivity distributions are presented for relatively able students, i.e. $a = a_r + \Delta a = 2.08$. Without on-the-job search the productivity density $h(p|s)$ is reduced to $f(p)/(1 - F(p_r(s)))$. This means that the equilibrium

productivity function is uniform truncated at the point $p_r(s)$. For the chosen scenario this reservation productivity is increasing over time, so the range of acceptable job offers is shrinking with higher schooling. For example $p_0(0) = 0.75$ for students with $s = 0$ and $p_0(T) = 0.85$ for most educated students. The fact that the distribution support is narrowing implies a positive correlation between productivity and schooling as more educated workers are sorted into more productive jobs. In the economy without on-the-job search the correlation coefficient⁴ between variables p and s is equal to 0.17.

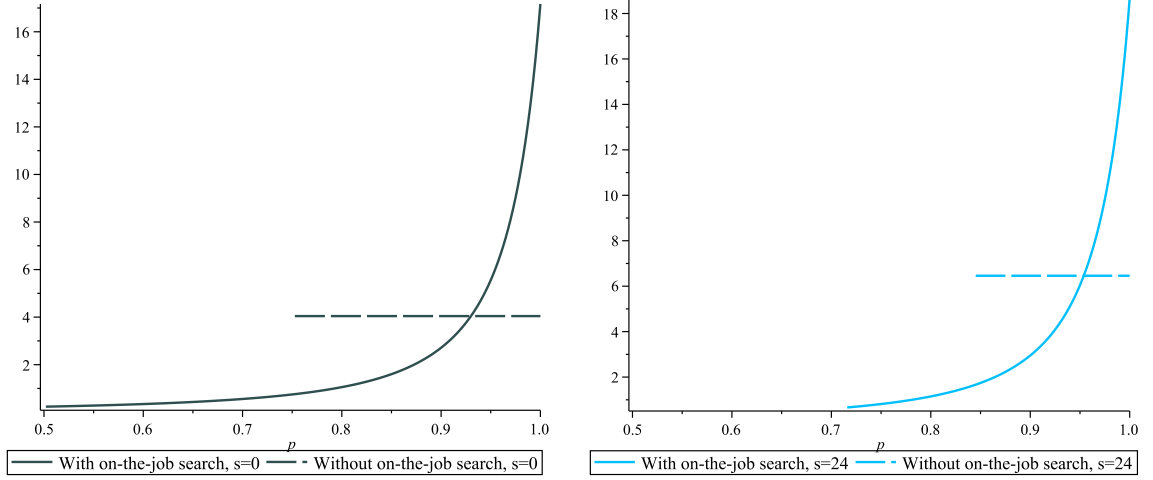


Figure 5: Conditional productivity density functions $h(p|s)$, $a=2.08$.

In the economy with a uniform distribution and on-the-job search the productivity density function $h(p|s)$ is increasing in p . This is because less (more) productive jobs are less (more) stable in the equilibrium. For the above mentioned scenario, when $a = 2.08$, the lowest acceptable productivity of unskilled workers is $p_r(0) = 0.50$ while the lowest productivity of the most educated workers is $p_r(T) = 0.72$. The correlation coefficient in the economy with on-the-job search is also positive as more educated workers select themselves into more productive jobs, nevertheless the correlation can be expected to be weaker as low productivity jobs are not long-lasting and constitute only a small proportion of the overall employment. In accordance with this reasoning the correlation coefficient⁵ between p and s in the economy with on-the-job search is equal to 0.06.

The first part of this section explains a twofold effect of on-the-job search on the average productivity of workers. First, there is a primary positive effect as changing jobs allows workers to climb a productivity ladder. However, there is also a negative countereffect as students initially accept jobs with lower productivity. In the numerical example it turns out that the two effects are mutually neutralizing each other and the average productivity remains virtually unchanged (0.88 vs. 0.91 for $s = 0$ and 0.92 vs. 0.93 for $s = T$).

Figure 6 illustrates two joint probability density functions $q(s, p)$ and $q_0(s, p)$ for the bench-

⁴The correlation coefficient is obtained as ($\mu_s = 8.44$, $\mu_p = 0.89$, $\sigma_s = 6.33$, $\sigma_p = 0.07$):

$$\text{corr}(p, s) = \frac{\text{cov}(p, s)}{\sigma_p \sigma_s}, \quad \text{where} \quad \text{cov}(p, s) = \int_0^T \int_{p_0(s)}^1 (s - \mu_s)(p - \mu_p)q(s, p)dpds$$

⁵The corresponding moments are: $\mu_s = 6.52$, $\mu_p = 0.88$, $\sigma_s = 5.88$ and $\sigma_p = 0.11$

mark scenario $a = a_r = 1.92$. The two bivariate density functions are given by:

$$q(s, p) = \frac{\rho f(p)(\rho + \lambda_e(1 - F(p_r)))g(s)}{(\rho + \lambda_e(1 - F(p_r)))^2(1 - F(p_r))} \quad \text{and} \quad q_0(s, p) = \frac{f(p)g_0(s)}{\int_0^T (1 - F(p_r(s)))g_0(s)ds}$$

with and without on-the-job search respectively. For a given productivity p both density functions are decreasing in s , which means that the proportion of employees with an above average schooling is lower than the proportion of employees with a below average schooling.

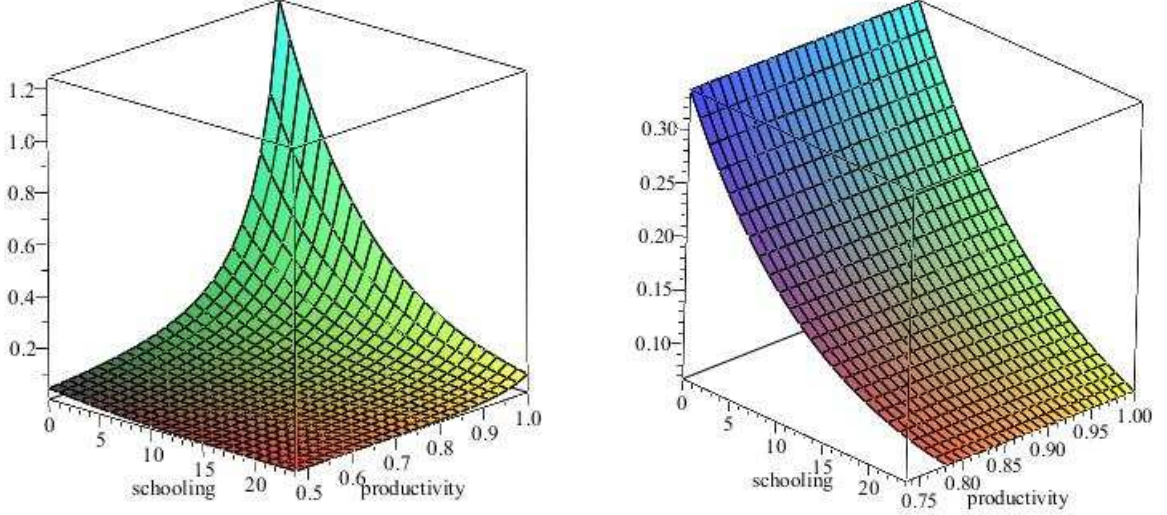


Figure 6: Left panel: with on-the-job search. Right panel: without on-the-job search

The situation is different when the schooling level is fixed. The joint density $q(s, p)$ is then increasing with a higher productivity p in the model with on-the-job search but it remains constant without it. A combination of these effects implies that low wages are paid to workers with low productivity and low schooling. However, for the case with on-the-job search figure 6 shows that the proportion of such jobs is relatively low. Similarly, there is a small proportion of workers employed at highest wages. Indeed, above average wages are paid to high-skilled workers employed in most productive jobs. In contrast, most workers in the economy are employed at moderate wages resulting from a combination of low schooling and reasonable match quality. In fact the mode of the distribution in the economy with on-the-job search is a point $p = \bar{p} = 1$ and $s = 0$ with a corresponding wage $w = \bar{p} = 1$. This explains why the equilibrium earnings distribution is pick-shaped with a unique interior mode on the distribution support (figure 7). This shape of the distribution is well-known in the empirical literature (Neal and Rosen (2000)).

Figure 7 presents the equilibrium distributions of wages in two economies with and without on-the-job search for the benchmark ability $a = a_r = 1.92$. The corresponding minimum wages are $p_r = 0.47$ and $p_0(0) = 0.73$, while the maximum wage is equal to $\bar{p}e^{\beta T} = 2.46$. Due to the decreasing schooling density both wage distributions are positively skewed in accordance with the empirical findings. The effect of on-the-job search is largely revealed in the left tail of the distribution. On the one hand, both the reservation productivity and the minimum wage are lower in the economy with on-the-job search. Therefore, wages in the range $[0.47, 0.73]$ are only

paid in the economy with on-the-job search. On the other hand, moderately paid jobs in the range $[0.77, 0.97]$ are more frequent in the economy without on-the-job search. This is explained by the inability of workers to change jobs in the hope of better payment. A combination of these impacts is illustrated on the right panel of figure 7.

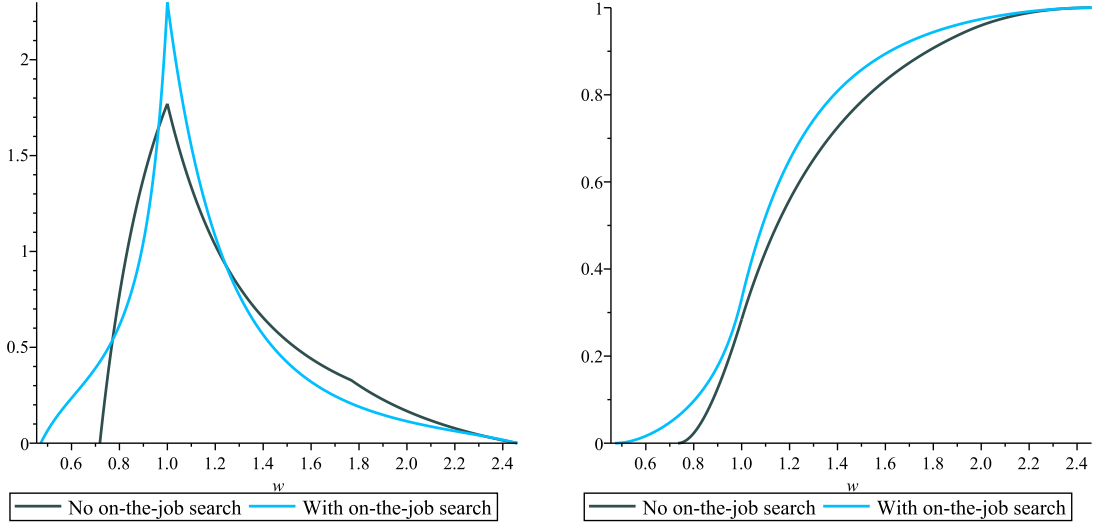


Figure 7: Left panel: probability density of w . Right panel: cumulative density of w

This figure shows that the equilibrium wage distribution without on-the-job search first-order stochastically dominates the alternative distribution: $Q(w) \geq Q_0(w) \forall w$. This result is already anticipated in the beginning of this section. Indeed, the average productivity of workers in the two economies is principally unchanged but the average schooling level is lower when the labour market is more flexible (8.9 vs. 7.0). These differences are translated into wages. In particular, the average wage is 1.16 in the economy with on-the-job search and it is 1.24 in the alternative economy. Therefore an early exit from the education system when the labour market is flexible may be associated with a severe wage penalty in the future.

5 Extension and calibration

This section extends the model and includes an additional state of unemployment. This extended model is then calibrated to match the standard values of the US labour market. In particular the model is extended in the following way. At every point in time students may fail to proceed to the next level in their education. The shock of exam failure arrives with a Poisson arrival rate γ . As an outcome of this shock the student enters the pool of unemployed workers and starts searching for a job. Intuitively, the γ -outflow of students from the education system is a proxy for the involuntary dropouts. The flow value of leisure when unemployed is denoted by b and the flow cost of active search is denoted by σ . Note that neither students nor employees search for jobs as a primary daily activity and so the cost of their search is assumed to be zero. Equation (3.1) is then modified as follows (for $z = y$):

$$kV(t) = \lambda \int_0^{\bar{p}} \max(W(t, p) - V(t), 0) dF(p) - \gamma(V(t) - U(t)) + \dot{V}(t)$$

where $U(t)$ is a present discounted value of unemployment with a schooling level t . The values of employment $W(s, p)$ and unemployment $U(s)$ are given by the following Bellman equations:

$$\begin{aligned} kU(s) &= b - \sigma + \lambda_u \int_0^{\bar{p}} \max(W(s, x) - U(s), 0) dF(x) \\ kW(s, p) &= pe^{\beta s} + \lambda_e \int_0^{\bar{p}} \max(W(s, x) - W(s, p), 0) dF(x) - \delta(W(s, p) - U(s)) \end{aligned}$$

where δ denotes the job destruction rate. Similar to the case of students let $b = \sigma$, this means that the flow utility of students and unemployed workers is normalized to zero. Unemployed workers follow the cutoff strategy such that $W(s, p_u) = U(s)$, which means that unemployed workers with schooling s accept the job if the match-specific productivity is above p_u and reject the job otherwise. The reservation productivity p_u is implicitly given by:

$$p_u = (\lambda_u - \lambda_e) \int_{p_u}^{\bar{p}} \frac{(1 - F(x))dx}{k + \delta + \lambda_e(1 - F(x))} \geq 0 \quad \text{if } \lambda_u \geq \lambda_e$$

and is independent of the schooling achievement s . This equation is due to the seminal contribution by Burdett and Mortensen (1998) and implies a positive relationship between the job-finding rate λ_u and the reservation productivity of unemployed workers.

Next consider the flows of workers in and out of the unemployment. Let $u(s)$ denote the measure of unemployed workers with a schooling level s . The differential equations for $u(s)$ and $n(s)$ are given by:

$$\dot{u}(s) = \gamma n(s) + \delta E(s) - \lambda_u(1 - F(p_u))u(s) - \rho u(s) \quad \dot{n}(s) = -(\rho + \gamma + \lambda(1 - F(p_r(s))))n(s)$$

where the first term $\gamma n(s)$ is an inflow of students unsuccessful in acquiring education, the second term $\delta E(s)$ is an inflow of workers losing jobs, while the last two terms together form an outflow from unemployment into employment and out of the labour force. Appendix VI shows that even with unemployment the reservation productivity of students $p_r(s)$ is a monotone function of time. Fix some particular level of schooling s , then either $p_r(s) \geq p_u$ or $p_r(s) < p_u$. Consider the former case, then the inflow of new hires from unemployment into $E(p|s)$ is equal to $\lambda_u(F(p) - F(p_u))u(s)$ for $p_u \leq p \leq p_r(s)$ and there is an additional inflow of students $\lambda(F(p) - F(p_r(s)))n(s)$ for $p_r(s) < p \leq \bar{p}$. Therefore the total measure of workers employed in jobs with a productivity below or equal to p and schooling s is given by:

$$E(p|s) = \begin{cases} \frac{\lambda_u(F(p) - F(p_u))u(s)}{\rho + \delta + \lambda_e(1 - F(p))} & \text{if } p_u \leq p \leq p_r(s) \\ \frac{\lambda(F(p) - F(p_r(s)))n(s)}{\rho + \delta + \lambda_e(1 - F(p))} + \frac{\lambda_u(F(p) - F(p_u))u(s)}{\rho + \delta + \lambda_e(1 - F(p))} & \text{if } p_r(s) < p \leq \bar{p} \end{cases}$$

where the outflow in the denominator $(\rho + \delta + \lambda_e(1 - F(p)))E(p|s)$ includes workers exiting the labour market, losing jobs or quitting voluntarily for more productive matches. Finally the joint probability density $q(s, p)$ can be found as $[\partial E(p|s)/\partial p]/E$, this derivation is presented in appendix VI. In the next section the model is calibrated to match the labour market in the US.

5.1 Calibration to the US labour market

Consider the model with unemployment and on-the-job search from the previous subsection. To match this model to the labour market indicators in the US let one period of time to be half a year (one semester). In addition let the maximum duration of schooling be equal to 12 years ($T=24$ semesters). This period includes the possibility to obtain a high school, bachelor, master and doctoral degrees. Further the annual interest rate is chosen to be 5% which means that $r = 0.025$. According to the US Bureau of Labor Statistics (BLS 2013), the total employment in May 2012 amounted to 130 million employees including 1 million high-school teachers, 1.5 million college teachers and 1.1 million scientists. So the total number of employees in research and teaching is equal to $l = 3.6$ million. In the steady state the inflow of workers into research and teaching $n(T)$ should be equal to the outflow ρl . To match parameters I consider the model with a constant reservation productivity of students p_r , the hazard rate in education is then given by $h \equiv \rho + \gamma + \lambda(1 - F(p_r))$, so the total number of high-skill workers becomes $n(T) = n(0) \cdot e^{-hT}$. This yields the first calibration equation between parameters $\rho, n(0)$ and h .

The total population in the model consists of employed and unemployed workers as well as students, where the total number of students is $n(0) \cdot (1 - e^{-hT})/h$. Targetting the unemployment rate of 6% which is a long-term average for the US gives rise to $8.3 = (130 \cdot 0.06)/(1 - 0.06)$ million unemployed individuals. Then the total population size can be written as $L = 138.3 + n(0) \cdot (1 - e^{-hT})/h$. Keeping the overall population at constant yields $n(0) = \rho L$ which is a second calibration equation between parameters $\rho, n(0)$ and h . Next, according to the report of the National Center for Education Statistics (2011) the total number of high-school students in the US is about 16.4 million. In the model this number can be generated as $n(0) \cdot (1 - e^{-8h})/h$ since the high-school duration is 4 years (8 months). This gives the third calibration equation:

$$n(0)e^{-24h} = 3.6\rho \quad n(0)(1 - e^{-8h})/h = 16.4 \quad n(0) = \rho(138.3 + n(0) \cdot (1 - e^{-24h})/h)$$

There is a unique solution vector to these equations $\{n(0) = 3.614, \rho = 0.022, h = 0.158\}$. This implies that the total number of students in the model is 22.3 million, while the total population is 160.6 million including 126.4 million employees in the regular labour market, 3.6 million workers in teaching and research, 8.3 million unemployed individuals and 22.3 million students. Moreover, the dropout rate γ can now be expressed as $\gamma = h - \rho - \lambda(1 - F(p_r)) = 0.136 - \lambda(1 - F(p_r))$, where both variables λ and p_r will be defined later in this section.

Following Jolivet et al. (2006) I set the annual job-finding rate of unemployed workers to 1.71 and the annual job destruction rate to 0.055, which means $\lambda_u = 0.86$ and $\delta = 0.0275$. Moreover, I choose the annual return to schooling to be 7.5% so that $\beta = 0.0375$ which is exactly in the middle of the range [5%..10%] suggested by Card (2001). Further, unemployment is stationary if $\gamma n + \delta E = (\rho + \lambda_u(1 - F(p_u)))u$, where $E = 126.4$ is the total employment in the regular market and δE is a flow of workers losing jobs per unit of time. With a uniform productivity distribution $F(x) = x, x \in [0..1]$ this implies:

$$\lambda_u(1 - p_u)u = \gamma n + \delta E - \rho u \quad \text{where} \quad \gamma = h - \rho - \lambda(1 - p_r) \quad (5.9)$$

Further assuming as before that the value of leisure is equal to the cost of active search ($b = \sigma$)

the reservation cutoffs of unemployed workers and students can be expressed as:

$$p_u = \left(1 - \frac{\lambda_e}{\lambda_u}\right)k\Omega(p_u) \quad p_r = \left(1 - \frac{\lambda_e(k + \gamma - \beta)}{\lambda k}\right)k\Omega(p_r) + \gamma U \frac{\lambda_e}{\lambda} \quad (5.10)$$

To avoid mass points in the earnings distribution λ is chosen so that $p_u = p_r$. Solving equations (5.9), (5.10) and $p_u = p_r$ produces the following values of the parameters $\gamma = 0.034$, $\lambda_e = 0.341$, $\lambda = 0.181$ as well as $p_r = p_u = 0.434$ (see table 1). These parameter choices imply $\lambda_u > \lambda_e > \lambda$. This ordering is intuitive as searching for jobs is the primary activity of the unemployed compared to students and employees. At the same time, in reality employed workers are more experienced and have larger social networks than students so the rate of job arrivals to students is the lowest. In addition, note that the $\rho + \lambda_u(1 - p_u) = 0.51$ so the average unemployment duration is equal to 1 year. As a robustness check I compare the average salary of workers involed in teaching and research with the average wage in the economy. In the real data (BLS 2013, annual wages in thousands) this ratio is equal to $\$67.60/\$45.79 = 1.5$ whereas in the model this ratio is $a_r/\bar{w} = 1.77/1.09 = 1.6$, where $a_r = 1.77$ corresponds to the stationary reservation productivity of students in the extended model with unemployment. This comparison reveals similarity of the observed ratio in the data with the one implied by the model.

Parameter	Value	Explanation
r	0.025	Annual interest rate 5%
ρ	0.022	Constant population size ($\rho = n(0)/L$)
γ	0.034	Proportion of teachers and researchers 2.2% (BLS 2013)
β	0.037	Net annual return to schooling 7.5% (Card (2001))
λ_u	0.860	Annual job-finding rate 1.710 (Jolivet et al. (2006))
δ	0.027	Annual job-destruction rate 0.055 (Jolivet et al. (2006))
λ_e	0.341	Unemployment rate 6%
λ	0.181	Target equal cut-off values $p_r = p_u$

Table 1: Parameter choices

The final wage density and the cdf in the labour market with endogenous schooling, on-the-job search and unemployment are illustrated in figure 8. In addition, some selected statistics of this distribution are presented in table 2. The skewness⁶ of the distribution is equal to 1.49, so the density function is skewed to the right. Also the median value (1.04) is lower than the average (1.09) which means that more than 50% of the employees are employed at wages below the average. Note that the unimodal shape of the distribution with a positive skewness is purely attributed to the endogenous factors of the model since the underlying uniform productivity distribution is symmetric. On the one hand the schooling density is downward-slopping as less and less workers achieve the higher levels of schooling and on the other hand at every level of schooling workers are moving upwards to more productive jobs. To see this consider two simplifications of the model. First, let jobs be homogeneous with a unique productivity value p which eliminates the gain from searching on-the-job. The pdf of the wage distribution is then:

$$q(w) = \frac{g\left(\frac{1}{\beta} \ln\left(\frac{w}{p}\right)\right)}{\beta w} = \frac{(\rho + \gamma + \lambda)(w/p)^{-(\rho+\gamma+\lambda)/\beta}}{\beta w(1 - e^{-(\rho+\gamma+\lambda)T})}$$

⁶The skewness is derived as ($\mu_w = 1.085$, $\sigma_w = 0.297$): $\int_{p_r}^{e^{\beta T}} \left(\frac{w - \mu_w}{\sigma_w}\right)^3 dQ(w)$

which is a strictly decreasing function of w . The first reason for the negative slope here is a decreasing density of the schooling distribution $g(s)$ and the second reason is that the wage is a convex function of schooling $w = pe^{\beta s}$ for a fixed productivity p .

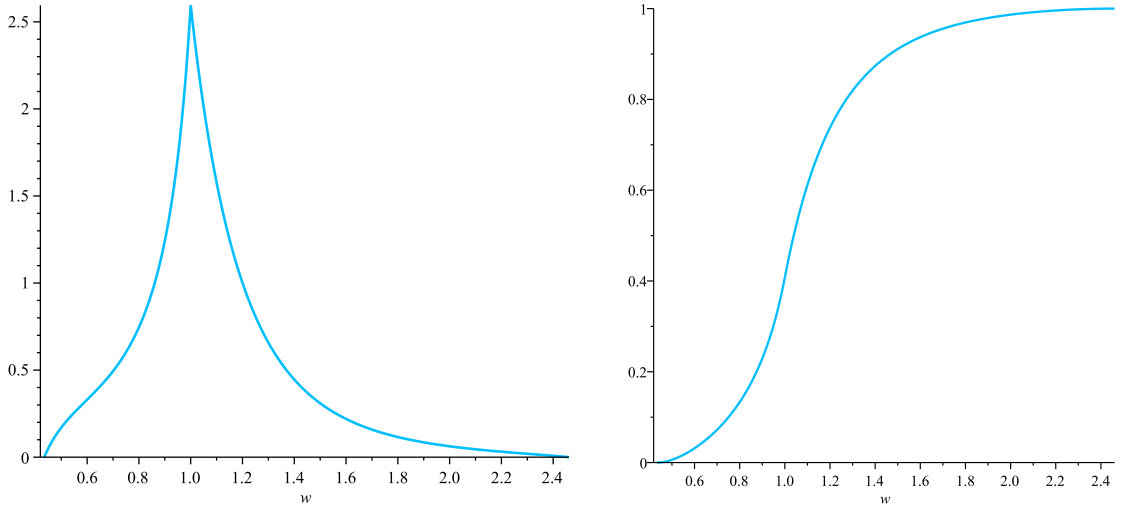


Figure 8: Left panel: probability density of w . Right panel: cumulative density of w

Minimum	Median	Mean	Maximum	Std. dev.	Skewness
0.434	1.039	1.085	2.461	0.297	1.488

Table 2: Selected statistics

Second, let workers be homogeneous with a unique schooling level s . The wage density is:

$$q(w) = e^{-\beta s} h(w e^{-\beta s}) = \frac{\rho e^{-\beta s} (\rho + \lambda_e (1 - p_u))}{(1 - p_u) (\rho + \lambda_e (1 - w e^{-\beta s}))}$$

and is a strictly increasing function of w since workers continue searching on-the-job. Thus heterogeneous workers and jobs are necessary to replicate the picked shape of the observed earnings distribution in the US (see figure 1). Finally note that the shape of the earnings density is principally different from the underlying uniform productivity distribution and it is the primary purpose of this paper to show that endogenous decisions of agents can generate a realistic earnings distribution. Nevertheless, an even better statistical fit of the density can be achieved by calibrating the model with a Pareto distribution of productivities suggested by Mortensen (2003). This will amplify the skewness and the range of earnings values.

6 Conclusions

This paper develops a matching model of the labour market with on-the-job search, heterogeneous jobs and endogenous schooling. Exit of students from the education system is modeled as an irreversible decision under uncertainty. This decision is fully characterised by the reservation productivity which is a monotone function of time. The reservation productivity is increasing over time in the case of high ability students. Intuitively, this means that high ability students become more choosy as their schooling level is increasing and accept a smaller range of employ-

ers. In the equilibrium this implies a positive sorting of high skill workers into more productive jobs. This finding is compatible with the empirical evidence if the productivity index in the model is associated with a firm size. The model then implies that larger more productive firms pay higher wages conditional on observed worker characteristics and employ a larger fraction of high skill workers.

Since the choice of schooling is made on the individual basis, the model predicts that workers are endogenously differentiated in the equilibrium with respect to the level of formal education. The schooling density is decreasing, so the proportion of workers is declining with a higher level of schooling. Intuitively, this means that there are less students pursuing a PhD degree than those graduating from a high school. Moreover, the average schooling level is lower in the economy with on-the-job search compared to the rigid labour market. This finding has important policy implications. It means that any policy intensifying on-the-job search, such a commuting subsidies, may have a negative indirect effect on the average schooling level in the population.

Finally, this paper investigates the shape of the equilibrium earnings density with a uniform productivity distribution. This functional choice guarantees that the shape of the equilibrium earnings density is not driven by the underlying assumption on the distribution of job types. For a given fixed level of schooling the model predicts that the earnings density is increasing. This effect is purely attributed to the process of on-the-job search as employees continue searching and change jobs in the search of better payment. Combined with an endogenous decreasing density of schooling types the model can generate an equilibrium earnings distribution with a long right tail and a unique mode on the distribution support. This means that the majority of workers earn wages in the middle range of the earnings distribution. At the same time there is a small proportion of employees in the beginning of their career with low wages and a small proportion of high-skill workers earning wages in the right tail of the distribution. The positive skewness of the earnings density is supported by empirical evidence.

7 Appendix

Appendix I. For simplicity consider first the case when $\rho = 0$ and let τ denote a spell of time between the two consequent offers. This means that τ is exponentially distributed with an intensity parameter λ . With a probability $P\{\tau \leq T - t\}$ the next job offer will arrive before the maximum education level T is reached. Let $E_\tau X(\tau)$ denote the expected present value of income of a student conditional on obtaining a job offer before T , so that:

$$\begin{aligned} V(t) &= P\{\tau \leq T - t\} E_\tau X(\tau) + P\{\tau > T - t\} V(T) e^{-r(T-t)} \\ &= \int_0^{T-t} X(\tau) \lambda e^{-\lambda\tau} d\tau + V(T) e^{-(r+\lambda)(T-t)} \end{aligned}$$

where $V(T)$ denotes the discounted present value of future income at time T .

$$\begin{aligned} X(\tau) &= \int_t^{t+\tau} (z - c) e^{-r(x-t)} dx + E_p[\max(W(t + \tau, p), V(t + \tau))] e^{-r\tau} \\ &= \frac{z - c}{r} (1 - e^{-r\tau}) + E_p[\max(W(t + \tau, p) - V(t + \tau), 0)] e^{-r\tau} + V(t + \tau) e^{-r\tau} \end{aligned}$$

$$= \frac{z-c}{r}(1-e^{-r\tau}) + \int_0^{\bar{p}} \max(W(t+\tau, p) - V(t+\tau), 0) dF(p) e^{-r\tau} + V(t+\tau) e^{-r\tau}$$

where $W(s, p)$ denotes the discounted present value of income of a worker employed with a productivity level p and the accumulated schooling time s .

$$\begin{aligned} V(t) &= (z-c) \frac{\lambda}{r} \int_0^{T-t} (1-e^{-r\tau}) e^{-\lambda\tau} d\tau + V(T) e^{-(r+\lambda)(T-t)} \\ &+ \lambda \int_0^{T-t} \left[\int_0^{\bar{p}} \max(W(t+\tau, p) - V(t+\tau), 0) dF(p) + V(t+\tau) \right] e^{-(r+\lambda)\tau} d\tau \\ V(t) &= \frac{(z-c)}{r} (1-e^{-\lambda(T-t)}) - \frac{(z-c)\lambda}{r(r+\lambda)} (1-e^{-(r+\lambda)(T-t)}) + V(T) e^{-(r+\lambda)(T-t)} \\ &+ \lambda \int_t^T \left[\int_0^{\bar{p}} \max(W(x, p) - V(x), 0) dF(p) + V(x) \right] e^{-(r+\lambda)(x-t)} dx \end{aligned}$$

by substitution $x = t + \tau$ and $dx = d\tau$. Denote the last term in the above equation by $A(t)$:

$$A(t) = V(t) - \frac{(z-c)}{r} (1-e^{-\lambda(T-t)}) + \frac{(z-c)\lambda}{r(r+\lambda)} (1-e^{-(r+\lambda)(T-t)}) - V(T) e^{-(r+\lambda)(T-t)}$$

Differentiate $V(t)$ with respect to t :

$$\begin{aligned} \dot{V}(t) &= -(z-c) \frac{\lambda}{r} e^{-\lambda(T-t)} + (z-c) \frac{\lambda}{r} e^{-(r+\lambda)(T-t)} + V(T) (r+\lambda) e^{-(r+\lambda)(T-t)} \\ &- \lambda \int_0^{\bar{p}} \max(W(t, p) - V(t), 0) dF(p) - \lambda V(t) + (r+\lambda) A(t) \\ &= -(z-c) \frac{\lambda}{r} e^{-\lambda(T-t)} + (z-c) \frac{\lambda}{r} e^{-(r+\lambda)(T-t)} + V(T) (r+\lambda) e^{-(r+\lambda)(T-t)} \\ &- \lambda \int_0^{\bar{p}} \max(W(t, p) - V(t), 0) dF(p) - \lambda V(t) - (r+\lambda) V(T) e^{-(r+\lambda)(T-t)} \\ &+ (r+\lambda) \left[V(t) - \frac{(z-c)}{r} (1-e^{-\lambda(T-t)}) + \frac{(z-c)\lambda}{r(r+\lambda)} (1-e^{-(r+\lambda)(T-t)}) \right] \\ &= -(z-c) (1-e^{-\lambda(T-t)}) + rV(t) - \lambda \int_0^{\bar{p}} \max(W(t, p) - V(t), 0) dF(p) \end{aligned}$$

Finally, this equation should be augmented with a term $\rho V(t)$ when the exit rate ρ is positive.

Appendix II. First subtract equation (3.4) from (3.5) to obtain:

$$\Omega'_p \dot{p}_r(t) = (r + \rho - \beta)(\Omega(p_r(t)) - \Omega(p_r)) - (\Phi(p_r(t)) - \Phi(p_r)) \quad (7.11)$$

Consider the first case $p_r < p_r(T)$. This case applies if $a > \Omega(p_r)(r + \rho)e^{-\beta T}$. Since $p_r(t)$ is a continuous function of time there exists a finite time spell $\Delta t > 0$ such that $p_r < p_r(T - \Delta t)$. Then equation (7.11) implies that $\dot{p}_r(T - \Delta t) > 0$ and therefore $p_r(T - \Delta t) < p_r(T)$, so that $p_r(t)$ is increasing in the time spell $[T - \Delta t, T]$. Repeat this step by induction for time spells $[T - (n-1)\Delta t, T - (n-2)\Delta t] \dots [T - 2\Delta t, T - \Delta t]$, where $2 < n < T/\Delta t$ takes positive integer values. Consider the last time spell $[T - (n-1)\Delta t, T - (n-2)\Delta t]$ such that $p_r < p_r(T - (n-1)\Delta t) < p_r(T)$. Next suppose that $p_r \geq p_r(T - n\Delta t)$. From equation (7.11) it follows that $\dot{p}_r(T - n\Delta t) \leq 0$ and therefore $p_r \geq p_r(T - n\Delta t) \geq p_r(T - (n-1)\Delta t)$. This is a

contradiction, so that $p_r(t) > p_r \forall t \in [0..T]$ and $\dot{p}_r(t) > 0$. In particular, from equation (3.6) this implies that $p_r(0) = p_r + c > p_r$, so that $c > 0$.

Similarly, for the opposite case $p_r > p_r(T)$ one obtains $p_r(t) < p_r \forall t \in [0..T]$ and $\dot{p}_r(t) < 0$, so that $p_r(0) = p_r + c < p_r$ and $c < 0$. This case applies if $a < \Omega(p_r)(r + \rho)e^{-\beta T}$.

Appendix III: The benchmark productivity p_r is obtained from $(k - \beta)\Omega(p_r) = \Phi(p_r)$, where variables $\Omega(p_r)$ and $\Phi(p_r)$ are given by:

$$k\Omega(p_r) = p_r + \lambda_e \int_{p_r}^{\bar{p}} \frac{(1 - F(x))dx}{k + \lambda_e(1 - F(x))} \quad \Phi(p_r) = \lambda \int_{p_r}^{\bar{p}} \frac{(1 - F(x))dx}{k + \lambda_e(1 - F(x))}$$

$$\begin{aligned} kd\Omega(p_r, \lambda_e) &= \frac{kdp_r}{k + \lambda_e(1 - F(p_r))} + kd\lambda_e \int_{p_r}^{\bar{p}} \frac{(1 - F(x))dx}{(k + \lambda_e(1 - F(x)))^2} \\ d\Phi(p_r, \lambda_e) &= \frac{-\lambda(1 - F(p_r))dp_r}{k + \lambda_e(1 - F(p_r))} - \lambda d\lambda_e \int_{p_r}^{\bar{p}} \frac{(1 - F(x))^2 dx}{(k + \lambda_e(1 - F(x)))^2} \end{aligned}$$

From $(k - \beta)d\Omega(p_r, \lambda_e) = d\Phi(p_r, \lambda_e)$ one obtains an expression for dp_r :

$$\frac{k - \beta + \lambda(1 - F(p_r))}{k + \lambda_e(1 - F(p_r))} dp_r = -d\lambda_e \int_{p_r}^{\bar{p}} \frac{(1 - F(x))(k - \beta + \lambda(1 - F(x)))}{(k + \lambda_e(1 - F(x)))^2} dx$$

Therefore, p_r is increasing in λ_e . Next insert dp_r into the equation for $kd\Omega(p_r, \lambda_e)$:

$$\begin{aligned} d\Omega(p_r, \lambda_e) &= \frac{-d\lambda_e}{k - \beta + \lambda(1 - F(p_r))} \int_{p_r}^{\bar{p}} \frac{(1 - F(x))(k - \beta + \lambda(1 - F(x)))}{(k + \lambda_e(1 - F(x)))^2} dx \\ &+ d\lambda_e \int_{p_r}^{\bar{p}} \frac{(1 - F(x))dx}{(k + \lambda_e(1 - F(x)))^2} = d\lambda_e \int_{p_r}^{\bar{p}} \frac{(1 - F(x))\lambda(F(x) - F(p_r))dx}{(k + \lambda_e(1 - F(x)))^2} \end{aligned}$$

So that $\Omega(p_r, \lambda_e)$ is decreasing with a lower λ_e despite a higher value of p_r .

Appendix IV:

$$\begin{aligned} k\Omega(p_r) &= p_r + \lambda_e \int_{p_r}^1 \frac{(1 - x)dx}{k + \lambda_e(1 - x)} = 1 - k \int_{p_r}^1 \frac{dx}{k + \lambda_e(1 - x)} \\ &= 1 + \frac{k}{\lambda_e} \ln(k + \lambda_e(1 - x)) \Big|_{p_r}^1 = 1 + \frac{k}{\lambda_e} \ln \frac{k}{k + \lambda_e(1 - p_r)} \\ \Phi(p_r) &= \lambda \int_{p_r}^1 \frac{1 - x}{k + \lambda_e(1 - x)} dx = \frac{\lambda}{\lambda_e} \left[1 - p_r + \frac{k}{\lambda_e} \ln \frac{k}{k + \lambda_e(1 - p_r)} \right] \end{aligned}$$

Without on-the-job search, the quadratic polynomial $(k - \beta)x - 0.5\lambda(1 - x)^2$ can be decomposed as $-0.5\lambda(x - p_1)(x - p_0)$. Integrating and using partial fractions one gets:

$$\begin{aligned} \int_{C_0}^{p_0(t)} \frac{dx}{-0.5\lambda(x - p_1)(x - p_0)} &= \frac{-1}{0.5\lambda(p_1 - p_0)} \int_{C_0}^{p_0(t)} \left[\frac{1}{x - p_1} - \frac{1}{x - p_0} \right] dx \\ &= \frac{-1}{0.5\lambda(p_1 - p_0)} \ln \frac{x - p_1}{x - p_0} \Big|_{C_0}^{p_0(t)} = \frac{-1}{0.5\lambda(p_1 - p_0)} \left[\ln \frac{p_0(t) - p_1}{p_0(t) - p_0} - \ln \frac{C_0 - p_1}{C_0 - p_0} \right] \end{aligned}$$

$$\begin{aligned}\ln \frac{p_0(t) - p_1}{p_0(t) - p_0} - \ln \frac{C_0 - p_1}{C_0 - p_0} &= -0.5\lambda(p_1 - p_0)t \\ p_0(t) \left[1 - \left(\frac{C_0 - p_1}{C_0 - p_0} \right) e^{-0.5\lambda(p_1 - p_0)t} \right] &= p_1 - p_0 \left(\frac{C_0 - p_1}{C_0 - p_0} \right) e^{-0.5\lambda(p_1 - p_0)t}\end{aligned}$$

$$p_0(t) = \frac{p_1(C_0 - p_0) - p_0(C_0 - p_1)e^{-0.5\lambda(p_1 - p_0)t}}{C_0 - p_0 - (C_0 - p_1)e^{-0.5\lambda(p_1 - p_0)t}}$$

$$\begin{aligned}\int_0^s p_0(t)dt &= \int_0^s \frac{p_1(C_0 - p_0) - p_0(C_0 - p_1)e^{-0.5\lambda(p_1 - p_0)t}}{C_0 - p_0 - (C_0 - p_1)e^{-0.5\lambda(p_1 - p_0)t}} dt \\ &= \frac{2}{\lambda} \left[0.5p_1\lambda t + \ln(C_0 - p_0 - (C_0 - p_1)e^{-0.5\lambda(p_1 - p_0)t}) \right] \Big|_0^s \\ &= \frac{2}{\lambda} \left[0.5p_1\lambda s + \ln(C_0 - p_0 - (C_0 - p_1)e^{-0.5\lambda(p_1 - p_0)s}) - \ln(p_1 - p_0) \right]\end{aligned}$$

$$\begin{aligned}\ln n(s) &= \ln n(0) - (\rho + \lambda)s + \lambda \int_0^s p_0(t)dt = \ln n(0) - (\rho + \lambda)s + p_1\lambda s \\ &+ 2 \ln(C_0 - p_0 - (C_0 - p_1)e^{-0.5\lambda(p_1 - p_0)s}) - 2 \ln(p_1 - p_0)\end{aligned}$$

$$n(s) = \frac{n(0)}{(p_1 - p_0)^2} \left[C_0 - p_0 - (C_0 - p_1)e^{-0.5\lambda(p_1 - p_0)s} \right]^2 e^{-(\rho + \lambda(1 - p_1))s}$$

Appendix V: This section relaxes the assumption of irreversibility when finishing education. Suppose employed workers may lose the job at rate δ . However, rather than staying unemployed they return back to the education system and continue studying. The present value equation of workers is then given by:

$$\begin{aligned}kW(s, p) &= pe^{\beta s} + \lambda_e \int_p^{\bar{p}} (1 - F(x)) W'_p dx - \delta(W(s, p) - V(s)) \\ W(s, p) &= \frac{ke^{\beta s}}{k + \delta} \underbrace{\left[\frac{p}{k} + \frac{\lambda_e}{k} \int_p^{\bar{p}} \frac{1 - F(x)}{k + \delta + \lambda_e(1 - F(x))} dx \right]}_{=\tilde{\Omega}(p)} + \frac{\delta}{k + \delta} V(s)\end{aligned}$$

Note that $\tilde{\Omega}(p)$ is an extended function with δ in the denominator, moreover, $\tilde{\Omega}(p)$ is a decreasing function of δ . Insert this expression into the present value equation for students:

$$\begin{aligned}kV(t) &= \lambda \int_0^{\bar{p}} \max \left[\frac{ke^{\beta t}}{k + \delta} \tilde{\Omega}(p) + \frac{\delta}{k + \delta} V(t) - V(t), 0 \right] dF(p) + \dot{V}(t) \\ kV(t)e^{-\beta t} &= \frac{\lambda k}{k + \delta} \int_0^{\bar{p}} \max [\tilde{\Omega}(p) - V(t)e^{-\beta t}, 0] dF(p) + \dot{V}(t)e^{-\beta t}\end{aligned}$$

Using notation $\phi(t) = V(t)e^{-\beta t}$ yields the following:

$$(k - \beta)\phi(t) = \underbrace{\frac{\lambda k}{k + \delta} \int_{p_r(t)}^{\bar{p}} (\tilde{\Omega}(p) - \phi(t)) dF(p)}_{=\tilde{\Phi}(p_r(t))} + \dot{\phi}(t)$$

The reservation productivity is defined as $\tilde{\Omega}(p_r(t)) = \phi(t)$, so that:

$$\tilde{\Omega}'(p_r(t))\dot{p}_r(t) = (k - \beta)\tilde{\Omega}(p_r(t)) - \tilde{\Phi}(p_r(t)) \quad \text{where} \quad \tilde{\Omega}'(p_r(t)) = \frac{k + \delta}{k(k + \delta + \lambda_e(1 - F(p_r(t))))}$$

Thus the reservation productivity $p_r(t)$ can be found from:

$$\int_C^{p_r(t)} \frac{\tilde{\Omega}'(x)dx}{(k - \beta)\tilde{\Omega}(x) - \tilde{\Phi}(x)} = t \quad \text{where} \quad \tilde{\Phi}(p_r(t)) = \frac{\lambda k}{k + \delta} \int_{p_r(t)}^{\bar{p}} (1 - F(p))\tilde{\Omega}'(p)dp$$

With a uniform productivity distribution I get that:

$$\begin{aligned} k\tilde{\Omega}(x) &= 1 + \frac{k + \delta}{\lambda_e} \ln \frac{k + \delta}{k + \delta + \lambda_e(1 - x)} \\ \tilde{\Phi}(x) &= \frac{\lambda}{\lambda_e} \left[1 - x + \frac{(k + \delta)}{\lambda_e} \ln \frac{k + \delta}{k + \delta + \lambda_e(1 - x)} \right] \end{aligned}$$

Note that for $\delta = 0$ both expressions simplify to yield $\Omega(x)$ and $\Phi(x)$. The end-point reservation productivity $p_r(T)$ can be found as before from the indifference condition $V(T) = a/k = \tilde{\Omega}(p_r(T))e^{\beta T}$. Then the change in $p_r(T)$ associated with a higher value of δ can be found as:

$$0 = \frac{\partial \tilde{\Omega}(p_r(T))}{\partial p_r(T)} dp_r(T) + \frac{\partial \tilde{\Omega}(p_r(T))}{\partial \delta} d\delta \Rightarrow \frac{\partial p_r(T)}{\partial \delta} = - \frac{\partial \tilde{\Omega}(p_r(T))/\partial \delta}{\partial \tilde{\Omega}(p_r(T))/\partial p_r(T)} > 0$$

This proves that a higher risk of losing the job δ , makes graduating students less willing to accept a job in the regular market, since the end-point reservation productivity $p_r(T)$ is rising. Intuitively, the present value of entering the regular market $\Omega(p_r(T))e^{\beta T}$ is decreasing with a higher job destruction rate δ , thus students require a higher productivity of the job in order to exit the education system. These effects are illustrated on figure 9 below.

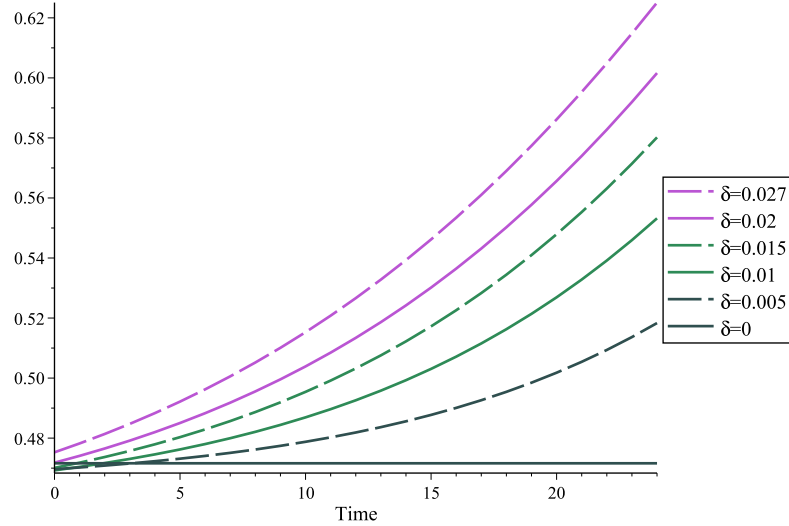


Figure 9: Effects of higher job destruction rate δ when schooling exit is reversible

Appendix VI

The present values of employment and unemployment can be expressed as:

$$kU(s) = \lambda_u \int_{p_u}^{\bar{p}} (1 - F(x)) W'_p dx \quad kW(s, p) = pe^{\beta s} + \lambda_e \int_p^{\bar{p}} (1 - F(x)) W'_p dx - \delta(W(s, p) - U(s))$$

where $W'_p = e^{\beta s} / (k + \delta + \lambda_e(1 - F(p))) = e^{\beta s} \Omega'(p)$. Multiply both sides of these equations by $e^{-\beta s}$ and substitute $U = U(s)e^{-\beta s}$ to obtain:

$$kU = \lambda_u \int_{p_u}^{\bar{p}} (1 - F(x)) \Omega'_p dx \quad k\Omega(p) = p + \lambda_e \int_p^{\bar{p}} (1 - F(x)) \Omega'_p dx - \delta(\Omega(p) - U)$$

The reservation productivity of unemployed workers is given by $U(s) = Ue^{\beta s} = \Omega(p_u)e^{\beta s} = W(s, p_u)$ which implies $U = \Omega(p_u)$. Inserting p_u into $\Omega(p)$ produces:

$$kU = \lambda_u \int_{p_u}^{\bar{p}} (1 - F(x)) \Omega'_p dx = p_u + \lambda_e \int_{p_u}^{\bar{p}} (1 - F(x)) \Omega'_p dx = k\Omega(p_u)$$

the present value equation for students is now modified as follows:

$$kV(t)e^{-\beta t} = \lambda \int_0^{\bar{p}} \max[\Omega(p) - V(t)e^{-\beta t}, 0] dF(p) + \dot{V}(t)e^{-\beta t} - \gamma(V(t) - U(t))e^{-\beta t}$$

As before substitute $V(t)e^{-\beta t} = \phi(t) = \Omega(p_r(t))$ to obtain:

$$(k + \gamma - \beta)\Omega(p_r(t)) = \lambda \int_{p_r(t)}^{\bar{p}} (1 - F(p)) \Omega'_p dp + \Omega'_p \dot{p}_r(t) + \gamma U = \Phi(p_r(t)) + \Omega'_p \dot{p}_r(t) + \gamma U$$

The reservation productivity of students $p_r(t)$ is then given by:

$$\int_C^{p_r(t)} \frac{\Omega'_p(x) dx}{(k + \gamma - \beta)\Omega(x) - \Phi(x) - \gamma U} = t, \quad t = 0..T$$

and satisfies the properties of proposition 1 where the stationary productivity p_r is now obtained from $(k + \gamma - \beta)\Omega(p_r) = \Phi(p_r) + \gamma U$.

Fix the level of schooling s and consider the case $p_r(s) \geq p_u$. The total employment of workers with a schooling level s can be obtained as $E(s) = E(\bar{p}|s)$:

$$(\delta + \rho)E(s) = \lambda(1 - F(p_r(s)))n(s) + \lambda_u(1 - F(p_u))u(s)$$

Insert the total measure of unemployed workers $u(s) = [\gamma n(s) + \delta E(s)] / [\rho + \lambda_u \bar{F}(p_u)]$ where $\bar{F}(x) = 1 - F(x)$ for the ease of exposition:

$$E(s) = \frac{[(\lambda \bar{F}(p_r(s)) + \gamma)(\rho + \lambda_u \bar{F}(p_u)) - \rho \gamma]}{\rho(\rho + \delta + \lambda_u \bar{F}(p_u))} n(s)$$

The total measure of workers with schooling s and productivity below or equal to p is given by:

$$E(p|s) = \frac{\lambda_u(F(p) - F(p_u))}{\rho + \delta + \lambda_e(1 - F(p))} \cdot \frac{\gamma n(s) + \delta E(s)}{\rho + \lambda_u(1 - F(p_u))} \quad \text{if } p_u \leq p \leq p_r(s)$$

$$E(p|s) = \frac{\lambda(F(p) - F(p_r(s)))n(s)}{\rho + \delta + \lambda_e(1 - F(p))} + \frac{\lambda_u(F(p) - F(p_u))}{\rho + \delta + \lambda_e(1 - F(p))} \cdot \frac{\gamma n(s) + \delta E(s)}{\rho + \lambda_u(1 - F(p_u))} \quad \text{if } p_r(s) \leq p < \bar{p}$$

8 References

ABOWD J., M., F. KRAMARZ, P. LENGERMANN, S. PEREZ-DUARTE (2004): "Are Good Workers Employed by Good Firms? A test of a simple assortative matching model for France and the United States", *CREST working paper*.

ALTONJI J. (1991): "The Demand for and Return to Education when Education Outcomes are Uncertain", *NBER Working paper Nr. 3714*.

ARAI M. (2003): "Wages, Profits, and Capital Intensity: Evidence from Matched Worker-Firm Data", *Journal of Labor Economics*, 21(3): 593-618.

AUD S., RAMANI A.K., FROHLICH L. (2011) "America's Youth: Transitions to Adulthood", *National Center for Education Statistics, Institute of Education Sciences*.

BAGGER J., F. FONTAINE, F. POSTEL-VINAY, J.-M. ROBIN (2011) "Tenure, Experience, Human capital and Wages: A Tractable Equilibrium Search Model of Wage Dynamics", *Working paper*.

BILKIC N., GRIES T., PILICHOWSKI (2012): "Stay in School or Start Working? - The Human capital Investment Decision under Uncertainty and Irreversibility", *Labour Economics*, 19(5): 706-717.

BURDETT K., C. CARILLO-TUDELA, M. COLES (2011): "Human Capital Accumulation and Labor Market Equilibrium", *International Economic Review*, 52(3): 657-677.

BURDETT K., D.T. MORTENSEN (1998): "Wage Differentials, Employer Size and Unemployment", *International Economic Review*, 39(2): 257-273.

CARD D. (1999): "The Causal Effect of Education on Earnings", *Handbook of Labor Economics*, edited by O. Ashenfelter and D. Card, Volume 3: 1801-1863.

CARD D. (2001): "Estimating the Return to Schooling: Progress on some Persistent Econometric Problems", *Econometrica*, 69(5): 1127-1160.

EATON J., H. S. ROSEN (1980): "Taxation, Human Capital, and Uncertainty", *The American Economic Review*, 70(4): 705-715.

GAUTIER P. (2002): "Unemployment and Search Externalities in a Model with Heterogeneous Jobs and Workers", *Economica*, 69: 21-40.

GROOT W., H. OOSTERBEEK (1992): "Optimal Investment in Human Capital under Uncertainty", *Economics of Education Review*, 11(1): 41-49.

HANCHANE S., A. LIOUI, D. TOUAHRI (2006): "Human Capital as a Risky Asset and the Effect of Uncertainty on the Decision to Invest", *Working paper, Aix-Marseille*.

HASKEL J., D. HAWKES, S. PEREIRA (2005): "Skills, Human Capital, and the Plant Productivity Gap: UK Evidence from Matched Plant, Worker and Workforce Data", *CEPR Discussion paper 5354*.

HOGAN V., I. WALKER (2007): "Education Choice under Uncertainty: Implications for Public Policy", *Labour Economics*, 14(6): 894-912.

IDSON T.L., D.J. FEASTER (1990): "A Selectivity Model of Employer-Size Wage Differentials", *Journal of Labor Economics*, 8(1): 99-122.

JACOBS B. (2007): "Real Options and Human Capital Investment", *Labour Economics*, 14(6): 913-925.

JOHNSON M. (2013): "Borrowing Constraints, College Enrollment, and Delayed Entry", *Journal of Labor Economics*, 31(4): 669-725.

JOLIVET G., POSTEL-VINAY F., ROBIN J. M. (2006): "The Empirical Content of the Job Search Model: Labour Mobility and Wage Distributions in Europe and the US", *European Economics Review*, 50: 877-907.

- JUDD K.L. (1998): "Taxes, Uncertainty, and Human Capital", *The American Economic Review*, 88(2): 289-292.
- LEVHARI D., Y. WEISS (1974): "The Effect of Risk on the Investment in Human Capital", *The American Economic Review*, 64(6): 950-963.
- MORTENSEN D.T. (2003): "Wage Dispersion: Why are Similar Workers Paid Differently?", *The MIT Press*.
- NAGYPAL E. (2008): "Worker Reallocation over the Business Cycle: the Importance of Employer-to Employer Transitions", *Working paper, Northwestern University*.
- NEAL D., S. ROSEN (2000): "Theories of the Distribution of Earning" in A. B. Atkinson and F.J. Bourguignon eds.*handbook of Income Distribution, Vol. 1*
- OI W.Y., T.L. IDSON (1999): "Firm Size and Wages", In O. Ashenfelter and D. Card eds.*Handbook of Labor Economics*, 3B: 2165-2214.
- OMMEREN J., G. VAN DEN BERG, C. GORTER (2000): "Estimating the Marginal Willingness to Pay for Commuting", *Journal of Regional Science*, 40(3): 541-563.
- PARK, J. H. (1994): "Returns to schooling: a peculiar deviation from linearity", Working paper no. 335 (Industrial Relations Section, Princeton University).
- ROWAN-KENYON H.T. (2007): "Predictors of Delayed College Enrollment and the Impact of Socioeconomic Status", *The Journal of Higher Education*, 78(2): 188-214.
- RUPERT P., E. STANCANELLI and E. WASMER (2009): "Commuting, Wages and Bargaining Power", *Annals of Economics and Statistics*, 95/96: 201-220.
- SHIMER R.(2012): "Reassessing the Ins and Outs of Unemployment", *Review of Economic Dynamics*, 15(2): 127-148.
- TROSKE K.R (1999) "Evidence on the Employer Size Wage Premium from Worker-Establishment Matched Data", *The Review of Economics and Statistics*, 81(1): 15-26.
- U.S. BUREAU of LABOR STATISTICS (2013): "Occupation Employment and Wages - May 2012".
- VAN DEN BERG G.J. (1990): "Non-stationarity in Job Search Theory", *The Review of Economic Studies*, 57(2): 255-277.
- WEINBERG D.H. (2004): "Evidence from Census 2000 about Earnings by Detailed Occupation for men and Women", *Census 2000 Special Reports*, US Bureau of Labor Statistics.
- WILLIAMS J.T. (1979): "Uncertainty and the Accumulation of Human Capital over the Life Cycle", *The Journal of Business*, 52(4): 521-548.
- YAMAGUCHI S. (2010): "Job Search, Bargaining and Wage Dynamics", *Journal of Labour Economics*, 28(3): 595-631.